

5 **CELLS HAVING AMPLIFIED SIGNAL TRANSDUCTION PATHWAY
 RESPONSES AND USES THEREFOR**

Cross Reference to Related Applications and Incorporation by Reference

 This application claims the benefit of U.S. Provisional Application No.
10 60/098,156, filed August 27, 1998, which is hereby incorporated by reference in its
 entirety. All patents, published patent applications and other references cited throughout
 this specification are hereby incorporated herein by reference in their entireties.

Background of the Invention

15 Cell surface receptors are an important class of proteins involved in cellular
 functioning because they are the primary mediators of cell to cell communication. For
 example, G protein coupled receptors (GPCRs) are an important category of cell surface
 receptors. The medical importance of these receptors is evidenced by the fact that more
 than 60% of all commercially available prescription drugs work by interacting with
20 known GPCRs.

 In their resting state, the G proteins, which consist of alpha (α), beta (β) and
 gamma (γ) subunits, are complexed with the nucleotide guanosine diphosphate (GDP)
 and are in contact with the receptors to which they are coupled. When a hormone or
 other first messenger binds to receptor, the receptor changes conformation and this alters
25 its interaction with the G protein. This causes the α subunit to release GDP, and the
 more abundant nucleotide guanosine triphosphate (GTP) displaces it, activating the G
 protein. The G protein then dissociates to separate the α subunit from the still
 complexed beta and gamma subunits. Either the $G\alpha$ subunit, or the $G\beta\gamma$ complex,
 depending on the pathway, interacts with an effector. The effector (which is often an
30 enzyme) in turn converts an inactive precursor molecule into an active "second
 messenger," which may diffuse through the cytoplasm or may be associated with
 downstream signal molecules, triggering a signal cascade. After a few seconds, the $G\alpha$
 converts the GTP to GDP, thereby becoming inactive. The inactivated $G\alpha$ may then
 reassociate with the $G\beta\gamma$ complex.

35 Hundreds, if not thousands, of receptors convey messages through heterotrimeric
 G proteins, of which at least 17 distinct forms have been isolated. Most G
 protein-coupled receptors are comprised of a single protein chain that is threaded
 through the plasma membrane seven times. Such receptors are often referred to as
 seven-transmembrane domain receptors (STRs). More than a hundred different GPCRs
40 have been found, including many distinct receptors that bind the same ligand, and there
 are likely many more GPCRs awaiting discovery.

5 The mating factor receptors of yeast cells (STE2 and STE3) also span the membrane of the yeast cell seven times and are coupled to yeast G proteins. Heterologous GPCRs can be expressed in yeast cells and can be made to couple to yeast G proteins resulting in the transduction of signals via the endogenous yeast pheromone system signaling pathway which is normally activated by STE2 or STE3. In some cases, 10 such heterologous receptors can be made to couple more effectively to the yeast pheromone system signaling pathway by coexpressing a heterologous G protein α subunit (e.g. U.S. Patent No. 5,482,835 of King et al), by expressing a chimeric G protein subunit (e.g. WO 94/23025), or by expressing a chimeric G protein coupled receptor (e.g., U.S. Patent No. 5,576,210 of Sledziewski et al.).

15 The $\beta\gamma$ subunits of the activated G protein stimulate the downstream elements of the pheromone system pathway, including the Ste20p protein kinase, and a set of kinases that are similar to MEK kinase, MEK (MAP kinase kinase), and MAP kinase of mammalian cells and are encoded by the STE11, STE7, and FUS3 genes, respectively (Whiteway et al. 1995. Science. 269:1572).

20 In recent years drug discovery has been advanced by expression of heterologous receptors in living cells. However, due to the complexities inherent in such heterologous expression studies, the development of reliable assays to search for modulators of these receptors has presented particular challenges. For example, it is often difficult to obtain sufficient signal amplification to enable efficient identification 25 of modulatory agents. The development of new drug discovery assays to identify novel modulators of receptors (e.g., GPCRs) would be of tremendous benefit.

Summary of the Invention

30 The present invention provides an important advance in drug screening methodologies for identifying modulators of receptors (e.g., G protein coupled receptors) by providing, *inter alia*, means by which ligand-induced signaling through a receptor can be amplified. In general, the invention involves amplifying ligand-generated signals by genetically engineering positive feedback loops at various levels of the signal transduction pathway. For example, a receptor-expressing cell can be 35 genetically engineered to carry a DNA construct in which a gene encoding a protein that activates the signal transduction pathway is operatively linked to a promoter that is responsive to activation of the signal transduction pathway. Thus, activation of the pathway by ligand binding to the receptor stimulates transcription of the activator protein, which in turn further stimulates activation of the pathway. Additionally or 40 alternatively, a negative regulator(s) of the pathway can be mutated to render it nonfunctional. Additionally or alternatively, an endogenous positive regulator of the

5 pathway be mutated to a supersensitive form, thereby increasing the sensitivity of the pathway to ligand stimulation.

In one embodiment, the invention pertains to a recombinant cell comprising:

- (i) a receptor that, upon ligand stimulation, activates an endogenous signal transduction pathway; and
- 10 (ii) a heterologous DNA construct comprising a gene encoding a protein that activates the signal transduction pathway, which gene is operatively linked to a promoter that is responsive to activation of the signal transduction pathway; wherein stimulation of the receptor by a ligand leads to expression of the heterologous DNA construct encoding the protein that activates the signal transduction pathway such
- 15 that signals generated by ligand binding to the receptor are amplified.

In a preferred embodiment, the cell is a yeast cell. In another embodiment, the cell is a mammalian cell. Preferred receptors include G protein coupled receptor, particularly heterologous G protein coupled receptors expressed by the recombinant cell. In one embodiment, the endogenous signal transduction pathway is a mitogen activated protein kinase (MAPK) pathway. In another embodiment, the endogenous signal

20 transduction pathway is a yeast pheromone response pathway.

The invention further provides a recombinant yeast cell comprising :

- (i) a heterologous G protein coupled receptor that, upon ligand stimulation, activates an endogenous yeast pheromone response pathway; and
- 25 (ii) a heterologous DNA construct comprising a gene encoding a protein that activates the yeast pheromone response pathway, which gene is operatively linked to a promoter that is responsive to activation of the yeast pheromone response pathway

wherein stimulation of the receptor by a ligand leads to expression of the heterologous DNA construct encoding the protein that activates the yeast pheromone response

30 pathway such that signals generated by ligand binding to the receptor are amplified.

Examples of proteins that activate the yeast pheromone response pathway and, accordingly, that can be used in the above recombinant cells, include STE5, STE4, STE12, STE11 and a dominant truncation allele of STE20. In one embodiment, a

35 hypersensitive mutant form of the protein that activates the yeast pheromone response pathway is used. For example, a hypersensitive mutant form of the Ste5 protein (e.g., having a T52M mutation or T52M/S18R mutation) can be used. Alternatively, a hypersensitive mutant form of the Ste4 protein (e.g., having a mutation selected from the group consisting of G124D, W136G, W136R, ΔL138, W136R/L138F and

40 W136G/S151C) can be used. Alternatively, a hypersensitive mutant form of the Ste11 protein (e.g., having a T596I mutation (allele Ste11-4) or a P278S mutation (allele

5 Ste11-1)) can be used. Alternatively, a hypersensitive mutant form of the Fus3 protein (e.g., having a I161L mutation) can be used.

The pheromone-responsive promoter that is used to drive expression of the protein that activates the pheromone response pathway can be, for example, a FUS1, AGA1, FAR1 or FUS2 promoter. A preferred promoter is the FUS1 promoter.

10 In another embodiment of the recombinant yeast cell, an endogenous yeast gene encoding a protein that negatively regulates the yeast pheromone system pathway is mutated to render the protein nonfunctional. For example, an endogenous gene encoding a phosphatase that negatively regulates the yeast pheromone system pathway can be mutated. Examples of endogenous phosphatases that can be mutated include
15 MSG5, PTP2, and PTP3.

In another embodiment, the cell further comprises a reporter gene construct that produces a detectable signal upon activation of the yeast pheromone response pathway.

In one embodiment, the heterologous DNA construct is carried by a high copy number plasmid. In another embodiment, the heterologous DNA construct is carried by
20 a low copy number plasmid.

Preferred yeast cells are *Saccharomyces cerevisiae* cells.

In a preferred embodiment, the invention provides a recombinant yeast cell comprising :

- (i) a heterologous G protein coupled receptor that, upon ligand stimulation,
25 activates an endogenous yeast pheromone response pathway; and
- (ii) a heterologous DNA construct comprising a STE5 or STE4 gene encoding a Ste5 or Ste4 protein, respectively, that activates the yeast pheromone response pathway, which gene is operatively linked to a FUS1 promoter that is responsive to activation of the yeast pheromone response pathway
30 wherein stimulation of the receptor by a ligand leads to expression of the heterologous DNA construct encoding the Ste5 or Ste4 protein that activates the yeast pheromone response pathway such that signals generated by ligand binding to the receptor are amplified.

In one embodiment, a hypersensitive mutant form of the Ste5 or Ste4 protein that
35 activates the yeast pheromone response pathway is used (e.g., mutations described above). In another embodiment, an endogenous yeast gene encoding a protein that negatively regulates the yeast pheromone system pathway is mutated to render the protein nonfunctional (e.g., a phosphatase, as described above).

Another aspect of the invention pertains to a recombinant yeast cell comprising:
40 a heterologous G protein coupled receptor that, upon ligand stimulation, activates an endogenous yeast pheromone response pathway,

5 wherein an endogenous yeast gene encoding a protein that negatively regulates the yeast pheromone system pathway is mutated to render the protein nonfunctional such that signals generated by ligand binding to the receptor are amplified.

Preferably, the endogenous gene that is mutated encodes a phosphatase that negatively regulates the yeast pheromone system pathway. Examples of such
10 phosphatases include MSG5, PTP2, and PTP3.

Another aspect of the invention pertains to a recombinant yeast cell comprising :
a heterologous G protein coupled receptor that, upon ligand stimulation,
activates an endogenous yeast pheromone response pathway,
wherein an endogenous FUS3 gene is mutated to encode a supersensitive form of Fus3
15 protein such that signals generated by ligand binding to the receptor are amplified.

Preferably, the supersensitive form of Fus3 protein has a I161L mutation.

The invention further encompasses use of the recombinant cells of the invention in screening assays to identify receptor modulators. For example, in one embodiment, the invention provides an assay to identify compounds that modulate the activity of a
20 receptor, comprising:

- (i) providing a recombinant cell of the invention that expresses a receptor, wherein a detectable signal is produced in the cell upon stimulation of the receptor;
- (ii) contacting the cell with a test compound; and
- (iii) identifying a compound which induces a change in the detectable signal in
25 the cell, such a change indicating that the compound modulates the activity of the receptor.

In one embodiment, the cell comprises a reporter gene construct that produces a detectable signal upon receptor stimulation.

30 ***Brief Description of the Drawings***

Figures 1A-1D are graphs that illustrate the effect of overexpression of STE5 in yeast cells expressing a heterologous C5a receptor FUS1-LacZ reporter gene transcription induced by treatment of the cells with the C5a receptor ligand ChaCha. β -
35 galactosidase activity (expressed as OD₅₇₅) (Figs. 1A and 1B) and induction ratios (relative to no ligand) (Figs. 1C and 1D) are plotted as a function of both ligand concentration (Figs. 1A and 1C) and log ligand concentration (Figs. 1B and 1D). The effect of overexpression of STE5 is compared to wildtype STE5.

40 Figure 2 is a schematic diagram which depicts yeast carrying a FUS1p-STE5 genetic construct in addition to the wild-type STE5. Initial activation of the pheromone response pathway occurs when ligand-stimulated receptor frees Ste4p/Ste18p, resulting

5 in the assembly of an activated MAPK-like pathway. Increased expression of
pheromone-responsive genes results in enhanced reporter gene expression and elevated
levels of Ste5p driven by the FUS1p-STE5 construct. Overexpressed Ste5p further
activates the pheromone response pathway resulting in amplified reporter gene
expression and Ste5p levels, thus leading to enhanced responsiveness to activation of the
10 pheromone pathway.

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted March 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5 *Detailed Description of the Invention*

 The present invention provides novel recombinant cells and assays utilizing such cells for screening and identifying pharmaceutically effective compounds that specifically modulate the activity of a receptor expressed by the cell (e.g., a heterologous G protein coupled receptor (GPCR) expressed the cell). The subject assays enable rapid
10 screening of large numbers of compounds (e.g., compounds in a library) to identify those which are receptor agonists or antagonists. Compositions of matter, such as novel recombinant cells (e.g., yeast cells) and novel gene constructs, are also embraced by the present invention. The instant assays provide a convenient format for discovering compounds which can be useful in modulating cellular function, as well as in
15 understanding the pharmacology of compounds that specifically interact with cellular receptors.

 The instant invention provides recombinant cells, and methods of use thereof, in which ligand stimulation of a receptor expressed by the cell leads to amplified signal transduction responses, through genetic manipulation of the cell, or components of the
20 signal transduction pathway thereof. The invention is based, at least in part, on the principle that yeast cells which express a genetic construct which comprises a pheromone responsive promoter linked to a gene encoding a protein which amplifies signaling via the pheromone response pathway display an amplification in signaling via the pheromone response pathway upon ligand stimulation through heterologous G
25 protein coupled receptors. Accordingly, this discovery has been used to develop yeast cells which express heterologous G protein coupled receptors and genetic constructs which amplify signaling via those receptors. The instant yeast cells display enhanced sensitivity to ligand induced stimulation of the heterologous receptor and therefore show improved properties in drug screening assays. In other embodiments, cells (e.g., yeast
30 cells) expressing a heterologous receptors are engineered such that a gene encoding a negative regulator of the signal transduction pathway is mutated to render the gene nonfunctional or, alternatively, a gene encoding a positive regulator of the pathway is mutated to render the regulator supersensitive (thereby increasing the sensitivity of ligand-induced responses).

35 Although the invention is described herein with particular reference to engineered yeast cells with amplified pheromone pathway responses, the ordinarily skilled artisan will readily appreciate that the approaches described herein for increasing the sensitivity of this yeast signal transduction pathway can be easily applied to other cell types and other signal transduction pathways. For example, the yeast pheromone
40 response pathway is analogous to the mitogen activated protein kinase (MAPK) pathway in mammalian cells and, accordingly, similar amplification schemes can be

5 accomplished in mammalian cells by manipulating the mammalian counterparts to the yeast pheromone system proteins in the same manner as the yeast proteins are manipulated as described herein.

In the practice of the instant invention, standard techniques known in the art can be used. See for example, Sherman. 1991. *Methods Enzymol.* 194:3; Sherman and
10 Hicks. 1991. *Methods Enzymol.* 194:21; Sambrook et al. *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989 or 1991 edition).

Before further description of the invention, certain terms employed in the specification, examples and appended claims are, for convenience, collected here.

15

9 I. Definitions

The term "compound" as used herein (e.g., as in "test compound") is meant to include both exogenously added test compounds and peptides endogenously expressed from a peptide library. For example, in certain embodiments, the reagent cell also
20 produces the test compound which is being screened. For instance, the reagent cell can produce. e.g., a test polypeptide, a test nucleic acid and/or a test carbohydrate which is screened for its ability to modulate the heterologous receptor activity. In such embodiments, a culture of such reagent cells will collectively provide a library of potential effector molecules and those members of the library which either agonize or
25 antagonize the receptor function can be selected and identified. Moreover, it will be apparent that the reagent cell can be used to detect agents which transduce a signal via the receptor of interest.

In other embodiments, the test compound is exogenously added. In such embodiments the test compound is contacted with the reagent cell. Exemplary
30 compounds which can be screened for activity include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, and natural product extract libraries. In such embodiments, both compounds which agonize or antagonize the receptor mediated signaling function can be selected and identified.

The term "non-peptidic compound" is intended to encompass compounds that are
35 comprised, at least in part, of molecular structures different from naturally-occurring L-amino acid residues linked by natural peptide bonds. However, "non-peptidic compounds" are intended to include compounds composed, in whole or in part, of peptidomimetic structures, such as D-amino acids, non-naturally-occurring L-amino acids, modified peptide backbones and the like, as well as compounds that are
40 composed, in whole or in part, of molecular structures unrelated to naturally-occurring

5 L-amino acid residues linked by natural peptide bonds. "Non-peptidic compounds" also are intended to include natural products.

As used herein, "recombinant cells" include any cells that have been modified by the introduction of heterologous DNA. Exemplary control cells include cells that are substantially identical to the recombinant cells, but do not express one or more of the
10 proteins encoded by the heterologous DNA, e.g., do not include or express a reporter gene construct, receptor or test polypeptide, or express a different heterologous DNA (e.g., a cell that expresses a different GPCR that couples to the same G protein as that of the GPCR whose activity is being examined).

As used herein, "heterologous DNA" or "heterologous nucleic acid" includes
15 DNA that does not occur naturally as part of the genome in which it is present or which is found in a location or locations in the genome that differs from that in which it occurs in nature or which is operatively linked to DNA to which it is not normally linked in nature (i.e., a gene that has been operatively linked to a heterologous promoter). Heterologous DNA is not naturally occurring in that position or is not endogenous to the
20 cell into which it is introduced, but has been obtained from another cell. Heterologous DNA can be from the same species or from a different species. In some embodiments, it is mammalian, e.g., human. Any DNA that one of skill in the art would recognize or consider as heterologous or foreign to the cell in which is expressed is herein encompassed by the term heterologous DNA. Examples of heterologous DNA include,
25 but are not limited to, genes which encode proteins that amplify signals transduced via the pheromone response pathway, DNA that encodes test polypeptides, receptors, reporter genes, transcriptional and translational regulatory sequences, or selectable or traceable marker proteins, such as a protein that confers drug resistance.

The terms "heterologous protein", "recombinant protein", and "exogenous
30 protein" are used interchangeably throughout the specification and refer to a polypeptide which is produced by recombinant DNA techniques, wherein generally, DNA encoding the polypeptide is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the heterologous protein. That is, the polypeptide is expressed from a heterologous nucleic acid.

As used herein, the term "genetic construct" refers to a heterologous gene (i.e.,
35 heterologous DNA or heterologous nucleic acid). which is operatively linked to a promoter which is responsive to signals transduced upon activation of the endogenous yeast pheromone signaling pathway. Since the promoter is responsive to activation of the signaling pathway, the heterologous gene is expressed upon activation of the
40 pathway. In preferred embodiments, the heterologous gene of the genetic construct is a gene which is of yeast origin and which activates the pheromone signaling pathway.

5 The expression of a genetic construct of the instant invention results in the amplification
of signals transduced by the activation of the pheromone response pathway.

As used herein, the term "high copy number plasmid" refers to a plasmid which exists in at least 5, or more, copies per cell, and more preferably in at least 10-20 copies per cell. The term "low copy number plasmid" refers to a plasmid which exists in fewer than 5 copies per cell, more preferably 2-3 copies, or less, per cell.

As used herein, the term "extracellular signal" is intended to encompass molecules and changes in the environment that are transduced intracellularly via cell surface proteins that interact, directly or indirectly, with the extracellular signal. An extracellular signal or effector molecule includes any compound or substance that in some manner alters the activity of a cell surface protein. Examples of such signals include, but are not limited to, molecules such as acetylcholine, growth factors and hormones, lipids, sugars and nucleotides that bind to cell surface receptors and modulate the activity of such receptors. The term, "extracellular signal" also includes as yet unidentified substances that modulate the activity of a cellular receptor, and thereby influence intracellular functions. Such extracellular signals are potential pharmacological agents that may be used to treat specific diseases by modulating the activity of specific cell surface receptors.

As used herein, "heterologous G protein receptor" is encoded by heterologous DNA and, upon expression of this heterologous DNA in a recombinant cell, the heterologous receptor is expressed in the recombinant cell. Exemplary GPCRs are described in detail herein.

The term "signal transduction" is intended to encompass the processing of physical or chemical signals from the extracellular environment through the cell membrane and into the cell, and may occur through one or more of several mechanisms, such as activation/inactivation of enzymes (such as proteases, or other enzymes which may alter phosphorylation patterns or other post-translational modifications), activation of ion channels or intracellular ion stores, effector enzyme activation via guanine nucleotide binding protein intermediates, formation of inositol phosphate, activation or inactivation of adenylyl cyclase, direct activation (or inhibition) of a transcriptional factor and/or activation. A "signaling pathway" refers to the components involved in "signal transduction" of a particular signal into a cell. The term "endogenous signaling pathway" indicates that some or all of the components of the signaling pathway are naturally-occurring components of the cell. An example of such a pathway is the endogenous pheromone system pathway of yeast.

40 The term "functionally couples to" (as in a receptor that is "functionally integrated into a signaling pathway in a cell" or "functionally integrated into an

5 endogenous yeast signaling pathway") is intended to refer to the ability of the receptor to be expressed at the surface of the cell and the ability of the expressed receptor to bind to modulators (e.g., a ligand of the receptor) and transduce signals into the cell via components of a signaling pathway of the cell. For example, a G protein coupled receptor (GPCR) which is functionally integrated into an endogenous pheromone
 10 response or signaling pathway of a yeast cell is expressed on the surface of the yeast cell, couples to a G protein of the pheromone response pathway within the yeast cell and transduces a signal in that yeast cell upon binding of a modulator to the receptor. For a G protein subunit to be functionally integrated into a yeast cell such a subunit, e.g., a chimeric, mutant or heterologous subunit, must be capable of coupling both to the
 15 GPCR and to the other G protein subunits, which can also be endogenous to the yeast cell, can be chimeric, or can be heterologous. A transduced signal may be detected by measuring any one of a number of responses to mating factors which occur in a yeast cell, e.g., growth arrest or transcription of an indicator gene responsive to signals produced by modulation of a pheromone system pathway.

20 The term "indicator gene" generically refers to an expressible (e.g., able to transcribed and (optionally) translated) DNA sequence which is expressed in response to a signal transduction pathway modulated by a target receptor or ion channel. Exemplary indicator genes include unmodified endogenous genes of the host cell, modified endogenous genes, or a reporter gene of a heterologous construct, e.g., as part of a
 25 reporter gene construct.

The term "endogenous gene" is intended to refer to a gene in a cell that is naturally part of the genome of the cell and which, most preferably, is present in its natural location in the genome (as opposed to "heterologous" DNA which has been introduced into the cell). Likewise, the term "endogenous protein" is intended to include
 30 proteins of a cell that are encoded by endogenous genes of the cell.

An endogenous gene that is to be used as an indicator gene may comprise the natural regulatory elements of the gene (e.g., the native promoter/enhancer elements that naturally regulate expression of the gene) or the endogenous gene can be "operatively linked to" (i.e., functionally coupled to) a "heterologous promoter" (or other
 35 heterologous regulatory elements). A "heterologous promoter" refers to a promoter that does not naturally regulate the gene to which the heterologous promoter is operatively linked. For example, an endogenous yeast gene that is not normally pheromone-responsive can be operatively linked to a heterologous promoter that is responsive to signals produced by the yeast pheromone system to thereby confer pheromone
 40 responsiveness on the endogenous yeast gene. Methods of using endogenous yeast

5 genes as indicator genes are described further in PCT Publication WO 98/13513, the contents of which are hereby expressly incorporated herein by this reference.

The term "detecting an alteration in a signal produced by an endogenous signaling pathway" (e.g., an endogenous yeast signaling pathway) is intended to encompass the detection of alterations in endogenous second messengers produced upon
10 activation of components of the endogenous signaling pathway, alterations in endogenous gene transcription induced upon activation of components of the endogenous signaling pathway, and/or alterations in the activity of an endogenous protein(s) upon activation of components of the endogenous signaling pathway. In certain embodiments, the term "detecting an alteration in a signal produced by an
15 endogenous signaling pathway" can also encompass assaying general, global changes to the cell such as changes in cell growth or cell morphology.

As used herein, a "reporter gene construct" refers to a nucleic acid that includes a "reporter gene" operatively linked to a transcriptional regulatory sequences. Transcription of the reporter gene is controlled by these sequences. The activity of at
20 least one or more of these control sequences is directly or indirectly regulated by the target receptor protein. The transcriptional regulatory sequences include the promoter and other regulatory regions, such as enhancer sequences, that modulate the activity of the promoter, or regulatory sequences that modulate the activity or efficiency of the RNA polymerase that recognizes the promoter, or regulatory sequences which are
25 recognized by effector molecules, including those that are specifically induced by interaction of an extracellular signal with the target receptor. For example, modulation of the activity of the promoter may be effected by altering the RNA polymerase binding to the promoter region, or, alternatively, by interfering with initiation of transcription or elongation of the mRNA. Such sequences are herein collectively referred to as
30 transcriptional regulatory elements or sequences. In addition, the construct may include sequences of nucleotides that alter translation of the resulting mRNA, thereby altering the amount of reporter gene product. The reporter gene constructs of the present invention provide a detectable readout in response to signals transduced in response to modulation of a heterologously expressed receptor.

35 The term "modulation", as in "modulation of a (heterologous) receptor" and "modulation of a signal transduction activity of a receptor protein" is intended to encompass, in its various grammatical forms, induction and/or potentiation, as well as inhibition and/or downregulation of receptor activity and/or one or more signal transduction pathways downstream of a receptor.

40 Agonists and antagonists are "receptor effector" molecules that modulate signal transduction via a receptor. Receptor effector molecules are capable of binding to the

5 receptor, though not necessarily at the binding site of the natural ligand or otherwise
modulating the activity of the receptor, for example, by influencing the activity of
components which regulate the receptor, or which function in the signal transduction
pathway initiated by the receptor. Receptor effectors can modulate signal transduction
when used alone, i.e. can be surrogate ligands, or can alter signal transduction in the
10 presence of the natural ligand or other known activators, either to enhance or inhibit
signaling by the natural ligand. For example, "antagonists" are molecules that block or
decrease the signal transduction activity of receptor, e.g., they can competitively,
noncompetitively, and/or allosterically inhibit signal transduction from the receptor,
whereas "agonists" potentiate, induce or otherwise enhance the signal transduction
15 activity of a receptor. The term "surrogate ligand" refers to an agonist which induces
signal transduction from a receptor.

The term "autocrine cell", as used herein, refers to a cell which produces a
substance which can stimulate a receptor located on or within the same cell as that which
produces the substance. For example, wild-type yeast MAT α and MATa cells are not
20 autocrine. However, a yeast cell which produces both α -factor and α -factor receptor, or
both a-factor and a-factor receptor, in functional form, is autocrine. By extension, cells
which produce a peptide which is being screened for the ability to activate a receptor
(e.g., by activating a G protein-coupled receptor) and also express the receptor are called
"autocrine cells". In some instances, such cells can also be referred to as "putative
25 autocrine cells" since some of the cells will express peptides from the library which will
not activate the receptor which is expressed. In a library of such cells, in which a
multitude of different peptides are produced, it is likely that one or more of the cells will
be "autocrine" in the stricter sense of the term.

As used herein, the term "not produced in functional form" with regard to
30 endogenous yeast proteins is intended to encompass proteins which are not produced in
functional form for any number of reasons, for example, because of a mutation to the
gene which encodes the protein or a deletion, e.g., a disruption, of the gene which
encodes the protein. The term "not produced in functional form" is also intended to
include conditional mutations (e.g. temperature sensitive mutation, wherein the protein
35 is not produced in functional form under certain conditions.

II. General Overview of Assay

As set out above, the present invention relates to yeast cell compositions and
methods for identifying effectors of a receptor protein or receptor protein complex. The
40 instant assays are characterized by the use of a mixture of recombinant yeast cells to
sample test compounds for receptor agonists or antagonists. As described in greater

5 detail below, the reagent cells express a heterologous GPCR protein functionally integrated into the cell and capable of transducing a detectable signal in the yeast cell. Exemplary GPCRs are discussed below. Compounds which either agonize or antagonize the receptor function can be selected and then identified based on biochemical signals produced by the receptor, or any more distal result of receptor-mediated stimulation, for example increases in endogenous mRNA expression, etc., or, 10 in some embodiments, by the use of reporter genes responsive to such signals. In certain embodiments, the library of compounds to be tested is a library of peptides which is expressed by the yeast cells and causes stimulation in an autocrine fashion.

The ability of compounds to modulate the signal transduction activity of the 15 target receptor can be scored for by detecting up or down-regulation of the detection signal. For example, GTPase activity, phospholipid hydrolysis, or protein phosphorylation stimulated by the receptor can be measured directly. Alternatively, the use of a reporter gene can provide a readout. In any event, a statistically significant change in the detection signal can be used to facilitate isolation of compounds of 20 interest.

As discussed in more detail below, the instant yeast cells are engineered to amplify signals transduced via the pheromone response pathway. In certain 25 embodiments, the yeast cells are modified to overexpress components of the pheromone response pathway which result in activation of the pathway. In other embodiments, the yeast cells express mutant forms of certain components of the pathway which lead to amplification of the response transduced by a heterologous receptor. In yet other embodiments, the instant yeast cells are modified to reduce the expression or activity of components of the pheromone response pathway which are responsible for negative regulation of the pathway.

30 In certain embodiments, the yeast cells for use in the instant assays express heterologous GPCR and an endogenous G protein subunit which couples to that receptor. Preferably, the yeast cells of the present invention have been modified such that coupling of the GPCR to the yeast pheromone signaling pathway is enhanced. For example, in preferred embodiments, the yeast cells express a heterologous GPCR and 35 mutated endogenous G protein subunit which facilitates functional integration of that receptor into the yeast cell. In another preferred embodiment, the yeast cells express a heterologous GPCR and a heterologous G protein subunit. In particularly preferred embodiments, the heterologous GPCR and the heterologous G protein subunit are of the same origin, e.g., mammalian. In yet another preferred embodiment, the yeast cells 40 express a mutated heterologous G protein subunit.

5 In still another preferred embodiment, the yeast cells express a chimeric G protein subunit. In particularly preferred embodiments the heterologous GPCR and the heterologous segment of the chimeric G protein subunit are derived from the same source. In more preferred embodiments, the second amino acid sequence in the G protein subunit chimera is derived from a mammalian G protein subunit. In particularly preferred embodiments, the second amino acid sequence is derived from a human G protein subunit sequence.

10 It will further be understood that the above embodiments are not mutually exclusive. For example, in certain preferred embodiments, a yeast cell may express a first mutated or chimeric G protein subunit and a second, different mutated or chimeric G protein subunit to enhance coupling to the heterologous receptor.

15 In certain embodiments the yeast cells also express an indicator gene that produces a detectable signal upon functional coupling of the heterologous G protein coupled receptor to the G protein. In certain embodiments the indicator gene is a reporter gene construct which including a reporter gene in operative linkage with one or more transcriptional regulatory elements responsive to the signal transduction activity of the target receptor, with the level of expression of the reporter gene providing the receptor-dependent detection signal. The amount of transcription from the reporter gene may be measured using any method known to those of skill in the art. For example, specific mRNA expression may be detected using Northern blots or specific protein product may be identified by a characteristic stain or an intrinsic activity. In preferred embodiments, the gene product of the reporter is detected by an intrinsic activity associated with that product. For instance, the reporter gene may encode a gene product that, by enzymatic activity, gives rise to a detection signal based on color, fluorescence, or luminescence.

20 The amount of activation of the indicator gene, e.g., expression of a reporter gene, is then compared to the amount of expression in either the same cell in the absence of the test compound or it may be compared with the amount of transcription in a substantially identical cell that lacks the specific receptors. A control cell may be derived from the same cells from which the recombinant cell was prepared but which had not been modified by introduction of heterologous DNA, e.g., the encoding a test polypeptide. Alternatively, it may be a cell in which the specific receptors are removed. Any difference, e.g., a statistically significant difference, in the amount of transcription indicates that the test compound has in some manner altered the activity of the specific receptor.

25 The amount of activation of the indicator gene, e.g., expression of a reporter gene, is then compared to the amount of expression in either the same cell in the absence of the test compound or it may be compared with the amount of transcription in a substantially identical cell that lacks the specific receptors. A control cell may be derived from the same cells from which the recombinant cell was prepared but which had not been modified by introduction of heterologous DNA, e.g., the encoding a test polypeptide. Alternatively, it may be a cell in which the specific receptors are removed. Any difference, e.g., a statistically significant difference, in the amount of transcription indicates that the test compound has in some manner altered the activity of the specific receptor.

30 In other preferred embodiments, the reporter gene provides a selection method such that cells in which the compound is an effector for the receptor have a growth

5 advantage. For example the reporter could enhance cell viability, relieve a cell nutritional requirement, and/or provide resistance to a drug.

By using any one of these readouts, compounds which modulate signaling via the heterologous receptor can be selected. If the compound does not appear to modulate signaling via the receptor protein, the assay may be repeated and modified by the
10 introduction of a step in which the recombinant cell is first contacted with a known activator of the target receptor to induce signal transduction from the receptor, and the compound is assayed for its ability to inhibit the activity of the receptor, e.g., to identify receptor antagonists. In yet other embodiments, compounds can be screened for members which potentiate the response to a known activator of the receptor.

15

III. Amplification of Signaling Via the Pheromone Response Pathway.

In general, the invention relates to amplification of ligand-generated signals by genetically engineering positive feedback loops at various levels of the signal transduction pathway. For example, a receptor-expressing cell can be genetically
20 engineered to carry a DNA construct in which a gene encoding a protein that activates the signal transduction pathway is operatively linked to a promoter that is responsive to activation of the signal transduction pathway. Thus, activation of the pathway by ligand binding to the receptor stimulates transcription of the activator protein, which in turn further stimulates activation of the pathway. Additionally or alternatively, a negative
25 regulator(s) of the pathway can be mutated to render it nonfunctional. Additionally or alternatively, an endogenous positive regulator of the pathway be mutated to a supersensitive form, thereby increasing the sensitivity of the pathway to ligand stimulation.

The invention generally provides compositions and methods that provide
30 amplified signaling via the pheromone response pathway through heterologous G protein coupled receptors (particularly, but not necessarily exclusively, since the triggering event may occur downstream of the GPCR) expressed in yeast cells. In certain embodiments, components of the yeast pheromone system which positively regulate signaling via the pathway are overexpressed under the control of a pheromone
35 responsive promoter. This strategy results in a system which is highly sensitive and responsive to activation of the pheromone system by heterologous GPCRs. In another embodiment, mutant forms of these components, which exhibit a hypersensitive phenotype (e.g., mutant forms that exhibit activity in the absence of activation by upstream components of the pheromone signaling pathway) are expressed in the yeast
40 host cells to amplify signaling via the pheromone pathway. In yet other embodiments,

5

In one aspect, the invention provides a recombinant cell comprising:

In another aspect, the invention provides a recombinant yeast cell comprising :

5 overexpressed in the yeast cells using a genetically engineered DNA construct is a dominant truncation allele of STE20 that is known in the art (described further in Ramer and Davis (1993) *Proc. Natl. Acad. Sci. USA* 90:452-456)

Exemplary pheromone responsive promoters which can be used to drive the expression of components which activate the yeast pheromone system pathway (such as
10 Ste5, Ste4, Ste11, Ste12 or the dominant truncation allele of Ste20) include the FUS1, AGA1, FAR1 or FUS2 promoter. A preferred promoter is the FUS1 promoter.

Accordingly, ligand-generated signals may be amplified by genetically engineering positive feedback loops at various levels of the pheromone response pathway. Strategies that use the FUS1 promoter to express Ste4p, for instance, establish
15 the positive feedback early in the pathway, before the MAPK module of Ste11p, Ste7p, and Fus3p. Strategies that use FUS1p-STE5, on the other hand, target the MAPK module itself, exploiting the ability of Ste5p to assemble the MAPK module and thereby positively influence the activity of the module. Conceptually similar approaches can be applied to events late in the pathway, beyond the MAPK module. For example, a
20 FUS1p-STE12 transcriptional unit may establish a positive feedback loop that amplifies Ste 12p-dependent readouts in response to ligand stimulation. Such an effect might be expected from the observation that overproduction of Ste12p leads to increased transcription of pheromone-inducible genes (Dolan and Fields, *Genes & Development* 4:492, 1990). Indeed, a positive feedback loop so far distal in the pheromone response
25 pathway may be particularly effective in amplifying ligand-generated signals, because it may avoid the action of such negative modulators as the phosphatases that down-regulate Fus3p.

In preferred embodiments, the genetic construct carrying the gene encoding the activator protein, operatively linked to the pheromone responsive promoter, is carried by
30 a plasmid. In certain instances, it may be desirable to control the copy number of the genetic construct. For example, in certain embodiments, expression of a gene which results in activation of the pheromone response pathway is from a low copy number in order to better detect ligand induced signaling. Exemplary low copy number plasmids suitable for use in yeast cells are known in the art and include, e.g., ARS vectors or centromeric sequences (CEN) (See e.g., Romanos et al. 1992. *Yeast* 8:423). In other
35 embodiments, however, the use of high copy number plasmids will be desirable. Exemplary high copy number plasmids are also known in the art and include *E. coli*-yeast shuttle vectors based on 2 μ . In yet other embodiments it may be desirable to express the genetic construct in a yeast cell using integrating vectors, such as YIp
40 vectors. Alternatively, DNA can be integrated into reiterated chromosomal DNA to generate stable multi-copy integrants (Kingsman et al. 1985. *Biotechnol. Genet. Eng. Revs.* 3:377; Lopes et al. 1989. *Gene* 79:199).

5 In another embodiment, a mutant form of a protein that activates the pheromone response pathway is used, which mutant form is expressed in the yeast host cell to allow for amplification of signalling by the pheromone pathway. These mutant forms include, for example, proteins that exhibit constitutive activity and/or proteins that have activity in the absence of activation by upstream components of the pheromone signaling
10 pathway. Mutant forms of proteins such as Ste5, Ste4, Ste11, Ste12 and Fus3 can be used.

A preferred mutant form is a mutant Ste5p containing the mutation T52M (*i.e.*, in which threonine at position 52 is replaced with methionine) or a double mutation of T52M and S18R (in which serine at position 18 is replaced with arginine) (referred to
15 herein as T52M,S18R). This mutant has been shown to have activity in the absence of activation by upstream components of the pheromone signaling pathway (Hansson et al. (1994) *Mol Cell Biol* 14:1054-1065). This mutant form of Ste5, referred to as STE5^{Hyp}, can be expressed under the control of a pheromone responsive promoter to allow for amplification of signalling by heterologously expressed GPCRs in the yeast
20 host cells.

In addition, dominant STE4 alleles have been identified which are constitutively activated for the pheromone response (Blinder et al. 1989. *Cell*. 56:479; Whiteway et al. 1994. *Mol. Cell. Biol.* 14:3223). Examples of preferred hypersensitive mutant forms of the Ste4 protein include those having a mutation selected from the group consisting of
25 G124D, W136G, W136R, ΔL138, and the double mutants W136R,L138F and W136G,S151C. These mutations are described in detail in Whiteway et al. 1994. *Mol. Cell. Biol.* 14:3223.

Constitutively active mutants of the STE11 gene have also been identified (Stevenson et al. 1992. *Genes and Dev.* 6:1293). For Ste11, constitutively active
30 mutations include substitution of an isoleucine for threonine at position 596 (T596I; for allele Ste11-4) and substitution of serine for proline at position 278 (P278S; for allele Ste11-1).

These constitutively active forms of STE4 and STE11 can also be expressed under the control of a pheromone responsive promoter to allow for amplification of
35 signalling by heterologously expressed GPCRs in the yeast host cells.

In yet other embodiments, yeast cells which express a mutant form of Fus3p can be used. For example, the I161L mutant of Fus3p (in which the isoleucine at position 161 has been substituted with leucine) has been shown to encode a MAPK that is hypersensitive to activity in the pheromone response pathway (*Mol Biol Cell* 5:297,
40 1994). Since this is a dominant allele, it can be used in strains that are wild-type for FUS3 and that carry a genetic construct having a pheromone responsive promoter linked to a gene that activates the pheromone response pathway.

In a particularly preferred embodiment, the invention provides a recombinant yeast cell comprising:

(i) a heterologous G protein coupled receptor that, upon ligand stimulation, activates an endogenous yeast pheromone response pathway; and

(ii) a heterologous DNA construct comprising a STE5 or STE4 gene encoding a Ste5 or Ste4 protein, respectively, that activates the yeast pheromone response pathway, which gene is operatively linked to a FUS1 promoter that is responsive to activation of the yeast pheromone response pathway

wherein stimulation of the receptor by a ligand leads to expression of the heterologous DNA construct encoding the Ste5 or Ste4 protein that activates the yeast pheromone response pathway such that signals generated by ligand binding to the receptor are amplified.

The Ste5 or Ste4 protein overexpressed in the cells via the pheromone-responsive promoter can be the wildtype protein or a hypersensitive mutant form of the protein, as described above.

In addition to overexpressing in the yeast cells a protein that activates the yeast pheromone system pathway, the yeast cells can be modified such that an endogenous yeast gene encoding a protein that negatively regulates the yeast pheromone system pathway is mutated to render the protein nonfunctional. Negative feedback inhibition of pheromone response pathway sensitivity is largely mediated through the actions of phosphatases acting on the yeast MAPK, Fus3p (Genes Devel 11:1690, 1997).

Therefore, in certain embodiments, the yeast cells of the instant invention are modified such that phosphatases which negatively regulate signaling via the pheromone response pathway are inhibited, e.g., by mutation or deletion. Components have been identified which negatively regulate the pathway. For example, Msg5p, a dual specificity phosphatase, has been shown to down-regulate the pheromone-activated MAPK pathway by dephosphorylating the FUS3p MAPK (Doi et al. 1994. EMBO 13:61). Another phosphatase gene, PTP3, has also been shown to negatively regulate the pathway (Zhan et. al. 1997. Genes and Development. 11:1690).

In an exemplary embodiment, the gene encoding the phosphatase MSG5 in a strain that carries an integrated FUS1p-STE5 in place of the wild-type STE5 can be deleted to hypersensitize the pathway to receptor activity. In certain embodiments, one or more other phosphatases acting on FUS3, e.g., PTP2 and PTP3 can be deleted. These embodiments result in cells that have very low noise in the absence of receptor activation due to low basal levels of Ste5p, yet will be capable of responding to agonist through positively reinforced expression of Ste5p.

In certain embodiments, the cell does not overexpress an activator protein but only has been modified to mutate a negative regulator of the signal transduction

5 pathway. Accordingly, in another aspect, the invention provides a recombinant yeast cell comprising:

a heterologous G protein coupled receptor that, upon ligand stimulation, activates an endogenous yeast pheromone response pathway, wherein an endogenous yeast gene encoding a protein that negatively regulates the yeast
10 pheromone system pathway (e.g., a phosphatase) is mutated to render the protein nonfunctional such that signals generated by ligand binding to the receptor are amplified.

In yet another embodiment, the yeast cell is modified such that an endogenous positive regulator of the pheromone response pathway be mutated to a supersensitive form, thereby increasing the sensitivity of the pathway to ligand stimulation. A
15 preferred positive regulator is Fus3. Accordingly, in still another aspect, the invention provides a recombinant yeast cell comprising :

a heterologous G protein coupled receptor that, upon ligand stimulation, activates an endogenous yeast pheromone response pathway, wherein an endogenous FUS3 gene is mutated to encode a supersensitive form of Fus3
20 protein such that signals generated by ligand binding to the receptor are amplified. A preferred supersensitive mutant form of Fus3 has a I161L mutation.

The cells of the invention can be used in screening assays to identify receptor modulators. Accordingly, another aspect of the invention pertains to assays to identify compounds that modulate the activity of a receptor. In one embodiment, the assay
25 involves

(i) providing a receptor-expressing recombinant cell of the invention, wherein a detectable signal is produced in the cell upon stimulation of the receptor;
(ii) contacting the cell with a test compound; and
(iii) identifying a compound which induces a change in the detectable signal in
30 the cell, such a change indicating that the compound modulates the activity of the receptor.

In one embodiment, the cell comprises a reporter gene construct that produces a detectable signal upon receptor stimulation. Use of the cells of the invention in screening assays, and various components and features thereof (e.g., test compounds,
35 detection of detectable signals, and the like) are described in further detail in the following subsections.

Standard molecular biology techniques can be used to create genetic constructs carrying a gene that activates the pheromone response pathway (e.g., Ste4, Ste5, Ste11, Ste12), operatively linked to a pheromone responsive promoter, including the mutant
40 forms discussed above. For exemplary descriptions of the preparation of such constructs, see Examples 1 and 2. Additionally, standard cell transfection techniques can be used to introduce such constructs into the yeast host cells. Still further standard

- 5 methods for deleting or specifically mutating endogenous yeast genes can be used to generate the deletion or point mutants discussed above.

IV. Host Cells

- 10 The host cells of the present invention may be of any species of yeast which are cultivable and in which an exogenous receptor can be made to engage the appropriate signal transduction machinery of the host cell. Exemplary species include *Kluyverei lactis*, *Schizosaccharomyces pombe*, and *Ustilago maydis*, with *Saccharomyces cerevisiae* being preferred. Other yeast which can be used in practicing the present invention are *Neurospora crassa*, *Aspergillus niger*, *Aspergillus nidulans*, *Pichia*
15 *pastoris*, *Candida tropicalis*, and *Hansenula polymorpha*. The term "yeast", as used herein, includes not only yeast in a strictly taxonomic sense, i.e., unicellular organisms, but also yeast-like multicellular fungi or filamentous fungi.

V. Expression Systems

- 20 In general, it will be desirable that an expression vector be capable of replication in the host cell. Heterologous DNA may be integrated into the host genome, and thereafter is replicated as a part of the chromosomal DNA, or it may be DNA which replicates autonomously, as in the case of a plasmid. In the latter case, the vector will include an origin of replication which is functional in the host. In the case of an
25 integrating vector, the vector may include sequences which facilitate integration, e.g., sequences homologous to host sequences, or encoding integrases.

- Appropriate cloning and expression vectors for use with bacterial, fungal, yeast, and mammalian cellular hosts are known in the art, and are described in, for example, Powels et al. (*Cloning Vectors: A Laboratory Manual*, Elsevier, New York, 1985).
30 Mammalian expression vectors may comprise non-transcribed elements such as an origin of replication, a suitable promoter and enhancer linked to the gene to be expressed, and other 5' or 3' flanking nontranscribed sequences, and 5' or 3' nontranslated sequences, such as necessary ribosome binding sites, a poly-adenylation site, splice donor and acceptor sites, and transcriptional termination sequences.

- 35 A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YEP51, YEP52 and pYES2 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach et al. (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p. 83, incorporated by reference herein). These vectors can replicate in
40 *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication

5 determinant of the yeast 2 micron plasmid. Moreover, yeast integrating plasmids, such as YIP5 or YRP17, can be used. For additional descriptions of suitable yeast vectors, see Christianson, et al. (1992) *Gene* 110:119-122, which describes yeast episomal plasmid (YE_p-type) vectors of the pRS420 series, which are multifunctional high copy number (about 20 copies/cell) shuttle vectors. See also Sikorski & Hieter (1989) 10 *Genetics* 122:19-27, which describes transplacement vectors of the pRS series (e.g., YC_p and YI_p). In addition, drug resistance markers such as ampicillin can be used. Suitable promoters for function in yeast include the promoters for metallothionein, 3-phosphoglycerate kinase (Hitzeman et al., *J. Biol. Chem.* 255, 2073 (1980) or other glycolytic enzymes (Hess et al., *J. Adv. Enzyme Req.* 7, 149 (1968); and Holland et al. 15 *Biochemistry* 17, 4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phospho-fructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phospho-glucose isomerase, and glucokinase. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., EPO 20 Publn. No. 73,657. Other promoters, which have the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the afore-mentioned metallothionein and glyceraldehyde-3-phosphate dehydrogenase, as well as enzymes responsible for maltose and galactose 25 utilization. Finally, promoters that are active in only one of the two haploid mating types may be appropriate in certain circumstances. Among these haploid-specific promoters, the pheromone promoters MF_{al} and MF_{α1} are of particular interest.

In certain embodiments, it will be desirable to control the copy number of the heterologous gene which is expressed. For example, in certain embodiments, expression 30 of a gene which results in activation of the pheromone response pathway will be from a low copy number vector in order to better detect ligand induced signaling. Exemplary low copy number plasmids suitable for use in yeast cells are known in the art and include, e.g., ARS vectors or centromeric sequences (CEN) (See e.g., Romanos et al. 1992. *Yeast* 8:423). In other embodiments, however, the use of high copy number 35 plasmids will be desirable. Exemplary high copy number plasmids are also known in the art and include *E. coli*-yeast shuttle vectors based on 2 μ . In yet other embodiments it may be desirable to express heterologous DNA in a yeast cell using integrating vectors, such as YI_p vectors. Alternatively, heterologous DNA can be integrated into reiterated chromosomal DNA to generate suitable multi-copy integrants (Kingsman et al. 1985. 40 *Biotechnol. Genet. Eng. Revs.* 3:377; Lopes et al. 1989. *Gene* 79:199)

5 VI. Receptors

Numerous different receptor types can be expressed in yeast cells for use in the instant invention.

The "heterologous receptors" of the present invention may be any G protein-coupled receptor which is exogenous to the cell which is to be genetically engineered for the purpose of the present invention. This receptor may be, for example, a plant or animal cell receptor. Screening for binding to plant cell receptors may be useful in the development of, e.g., herbicides. In the case of an animal receptor, it may be of invertebrate or vertebrate origin. An invertebrate receptor would, for example, facilitate development of insecticides. The expression of a receptor from a different species of yeast is also included within the term "heterologous" and could be used in the development of fungicides. The receptor may also be a vertebrate, more preferably a mammalian, still more preferably a human, receptor. The exogenous receptor is also preferably a seven transmembrane segment receptor.

Known ligands for GPCRs include: purines and nucleotides, such as adenosine, cAMP, ATP, UTP, ADP, melatonin and the like; biogenic amines (and related natural ligands), such as 5-hydroxytryptamine, acetylcholine, dopamine, adrenaline, histamine, noradrenaline, tyramine/octopamine and other related compounds; peptides such as adrenocorticotrophic hormone (acth), melanocyte stimulating hormone (msh), melanocortins, neurotensin (nt), bombesin and related peptides, endothelins, cholecystokinin, gastrin, neurokinin b (nk3), invertebrate tachykinin-like peptides, substance k (nk2), substance p (nk1), neuropeptide y (npy), thyrotropin releasing-factor (trf), bradykinin, angiotensin ii, beta-endorphin, c5a anaphalatoxin, calcitonin, chemokines (also called intercrines), corticotrophic releasing factor (crf), dynorphin, endorphin, fmlp and other formylated peptides, follitropin (fsh), fungal mating pheromones, galanin, gastric inhibitory polypeptide receptor (gip), glucagon-like peptides (glps), glucagon, gonadotropin releasing hormone (gnrh), growth hormone releasing hormone(ghrh), insect diuretic hormone, interleukin-8, leutropin (lh/hcg), met-enkephalin, opioid peptides, oxytocin, parathyroid hormone (pth) and pthrp, pituitary adenylyl cyclase activating peptide (pacap), secretin, somatostatin, thrombin, thyrotropin (tsh), vasoactive intestinal peptide (vip), vasopressin, vasotocin; eicosanoids such as ip-prostacyclin, pg-prostaglandins, tx-thromboxanes; retinal based compounds such as vertebrate 11-cis retinal, invertebrate 11-cis retinal and other related compounds; lipids and lipid-based compounds such as cannabinoids, anandamide, lysophosphatidic acid, platelet activating factor, leukotrienes and the like; excitatory amino acids and ions such as calcium ions and glutamate.

5 Preferred G protein coupled receptors include: α 1A-adrenergic receptor, α 1B-adrenergic receptor, α 2-adrenergic receptor, α 2B-adrenergic receptor, β 1-adrenergic receptor, β 2-adrenergic receptor, β 3-adrenergic receptor, m1 acetylcholine receptor (AChR), m2 AChR, m3 AChR, m4 AChR, m5 AChR, D1 dopamine receptor, D2 dopamine receptor, D3 dopamine receptor, D4 dopamine receptor, D5 dopamine
 10 receptor, A1 adenosine receptor, A2b adenosine receptor, 5-HT1a receptor, 5-HT1b receptor, 5HT1-like receptor, 5-HT1d receptor, 5HT1d-like receptor, 5HT1d beta receptor, substance K (neurokinin A) receptor, fMLP receptor, fMLP-like receptor, angiotensin II type 1 receptor, endothelin ETA receptor, endothelin ETB receptor, thrombin receptor, growth hormone-releasing hormone (GHRH) receptor, vasoactive
 15 intestinal peptide receptor, oxytocin receptor, somatostatin SSTR1 and SSTR2, SSTR3, cannabinoid receptor, follicle stimulating hormone (FSH) receptor, leutropin (LH/HCG) receptor, thyroid stimulating hormone (TSH) receptor, thromboxane A2 receptor, platelet-activating factor (PAF) receptor, C5a anaphylatoxin receptor, Interleukin 8 (IL-8) IL-8RA, IL-8RB, Delta Opioid receptor, Kappa Opioid receptor, mip-1/RANTES
 20 receptor, Rhodopsin, Red opsin, Green opsin, Blue opsin, metabotropic glutamate mGluR1-6, histamine H2 receptor, ATP receptor, neuropeptide Y receptor, amyloid protein precursor receptor, insulin-like growth factor II receptor, bradykinin receptor, gonadotropin-releasing hormone receptor, cholecystokinin receptor, melanocyte stimulating hormone receptor, antidiuretic hormone receptor, glucagon receptor, and
 25 adrenocorticotrophic hormone II receptor.

Other suitable receptors are known in the art. The term "receptor," as used herein, encompasses both naturally occurring and mutant receptors.

Many of these G protein-coupled receptors, like the yeast a- and α -factor receptors, contain seven hydrophobic amino acid-rich regions which are assumed to lie
 30 within the plasma membrane. Thus, for expression in yeast, the gene could be operably linked to a promoter functional in the cell to be engineered and to a signal sequence that also functions in the cell. For example, suitable promoters include Ste2, Ste3 and gal10. Optionally, the codons of the gene would be optimized for expression in yeast. See Hoekema et al., (1987) *Mol. Cell. Biol.*, 7:2914-24; Sharp, et al., (1986) 14:5125-43.

35 In some instances a foreign receptor which is expressed in yeast will functionally integrate into the yeast membrane, and there interact with the endogenous yeast G protein. In other instances, either the receptor may be modified or a compatible G protein or a chimeric (i.e., part yeast/part mammalian) G protein subunit which can properly interact with the exogenous receptor G protein may be provided. The
 40 homology of STRs is discussed in Dohlman et al., *Ann. Rev. Biochem.*, (1991) 60:653-88. When STRs are compared, a distinct spatial pattern of homology is discernible. The

transmembrane domains are often the most similar, whereas the N- and C-terminal regions, and the cytoplasmic loop connecting transmembrane segments V and VI are more divergent. The functional significance of different STR regions has been studied by introducing point mutations (both substitutions and deletions) and by constructing chimeras of different but related STRs. Synthetic peptides corresponding to individual segments have also been tested for activity. Affinity labeling has been used to identify ligand binding sites. Such information can be useful in creating mutations in GPCRs to enhance functionality.

If a naturally occurring exogenous GPCR cannot be made functional in yeast, it may be mutated for this purpose. A comparison would be made of the amino acid sequences of the exogenous receptor and of the yeast receptors, and regions of high and low homology identified. Trial mutations would then be made to distinguish regions involved in ligand or G protein binding, from those necessary for functional integration in the membrane. The exogenous receptor would then be mutated in the latter region to more closely resemble the yeast receptor, until functional integration was achieved. If this were insufficient to achieve functionality, mutations would next be made in the regions involved in G protein binding. Alternatively, random mutagenesis can be performed, followed by selection for function in yeast.

Preferably, the yeast genome is modified so that it is unable to produce the yeast receptors which are homologous to the exogenous receptors in functional form in order to facilitate assay interpretation. For example, the endogenous G protein or G protein subunit is mutated generating, for example, a temperature sensitive mutant.

GPCR Expression

In other embodiments, a secretory signal of a yeast protein can be used to direct transport of the receptors, for example G protein coupled receptors, to the plasma membrane as described in the appended examples. Previous work has demonstrated the secretory expression of foreign proteins in yeast cells using the signal sequence of a yeast secreted protein such as invertase or acid phosphatase, encoded by the *SUC2* and *PHO5* genes, respectively (Schraber, M.D. et al. (1986) *Methods Enzymol.* 119:416; Moir, D.T. et al. (1991) *Methods Enzymol.* 194:491-507). The vast majority of the secreted proteins possess a hydrophobic N-terminal signal sequence which targets them to the endoplasmic reticulum. A leader sequence of the α -factor precursor encoded by the *MF α 1* gene was shown to promote the most efficient secretion of various heterologous proteins. In addition to a signal sequence, the α -factor leader includes a hydrophilic pro-region which is believed to facilitate protein transport at the later stages of the secretory pathway.

5 Both secreted and membrane proteins including G protein coupled receptors are delivered to the cell surface through the same secretory pathway. Some receptors, for example, metabotropic glutamate receptors and vasoactive intestinal peptide receptors, also possess the N-terminal signal sequence, whereas some do not. In the latter case, a first transmembrane domain is believed to interact with the ER translocation machinery.

10 The use of yeast secretory signals, in particular, the α -factor leader, may be desirable to provide the more efficient integration of the receptors into the membrane of the endoplasmic reticulum and transport to the plasma membrane. In fact, the cell surface expression of the rat M5 receptor directed by the α -factor leader has been documented (Huang et al. (1992) *Biochem. Biophys. Res. Commun.* 181:1180).

15

VII. G protein subunits and complexes

In certain instances it will be desirable to modify naturally occurring forms of yeast or mammalian G-protein subunits. For instance, where a heterologous GPCR does not adequately couple to the endogenous yeast G protein subunit, such a subunit, e.g.,

20 GPA1 may be modified to improve coupling. Such modifications can be made by mutation, e.g., directed mutation or random mutation, using methods known in the art and described in more detail below.

Alternatively, a heterologous subunit can be expressed. The specificity of coupling of a receptor to a heterotrimeric G-protein is largely determined by the α

25 subunit of the G-protein. Thus, in preferred embodiments, a heterologous $G\alpha$ subunit is expressed in the yeast cell. The predominant role of the yeast $G\alpha$, GPA1, is to bind to and sequester the effector-signaling $\beta\gamma$ component of the heterotrimer. Thus, in order to achieve functional integration into a yeast pheromone signaling pathway, a heterologous $G\alpha$ subunit must bind to yeast $\beta\gamma$ in the quiescent state, and release it upon receptor

30 activation.

If functional integration is not achieved, or is not optimal, the heterologous subunit can be mutated. For example, in general, mammalian $G\alpha$ subunits couple poorly to the $\beta\gamma$ subunits of yeast cells. In yeast which lack their own endogenous $G\alpha$ subunit, this failure to couple results in the constitutive activation of the pheromone

35 pathway due to the effector activity of the unbound yeast $\beta\gamma$. Accordingly, if a naturally occurring heterologous G protein subunit does not enhance coupling, modifications can be made. Such modifications may take the form of mutations which are designed to increase the resemblance of the G protein subunit to the yeast G protein subunit while decreasing its resemblance to the heterologous receptor-associated G protein subunit.

40 For example, a residue may be changed so as to become identical to the corresponding yeast G protein residue, or to belong to the same exchange group of that

5 residue. After modification, the modified G protein subunit might or might not be "substantially homologous" to the heterologous and/or the yeast G protein subunit.

In the case of $G\alpha$, modifications are preferably concentrated in regions of the $G\alpha$ which are likely to be involved in $G\beta\gamma$ binding.

In other embodiments, modifications will take the form of replacing one or more amino acids of the receptor-associated G protein subunit with the corresponding yeast G protein subunit amino acids, thereby forming a chimeric G protein subunit. In preferred embodiments, three or more consecutive amino acids are replaced. In other embodiments, point mutations may be sufficient.

Chimeric G protein subunits of the invention enhance coupling of the
15 heterologous receptor to the endogenous yeast signaling pathway. For example, a
chimeric G α subunit will interact with the heterologous receptor and the yeast G $\beta\gamma$
complex, thereby permitting signal transduction.

A yeast cell of the present invention can express one or more of the indicated G protein structures. For example, a yeast cell can express a chimeric G α subunit, and an endogenous yeast G $\beta\gamma$, a mammalian G $\beta\gamma$, a mutated mammalian G $\beta\gamma$, or a chimeric G $\beta\gamma$.

In preferred embodiments, both the receptor and the heterologous subunit are derived from the same source, e.g., are mammalian. In particularly preferred embodiment, both are human in origin.

25 In another preferred embodiment, a yeast cell that expresses a heterologous or
chimeric G protein subunit has been modified such that the endogenous, homologous G
protein subunit gene is disrupted.

In certain embodiments, yeast strains lacking pheromone receptors and having no heterologous receptor capable of coupling to the pheromone response pathway may be used to assess the affinity of an endogenous yeast G protein subunit, a mammalian G protein subunit, a mutated G protein subunit, or chimeric G protein subunit for other yeast subunits. For example, the affinity of *gpa1p*, chimeric *gpa*-G α s, or other G α subunit for yeast $\beta\gamma$ or other chimeric $\beta\gamma$ subunit can be assessed. Such strains depend on free $\beta\gamma$ for signaling through the pheromone response pathway leading to growth arrest. Mutant G α subunits may be tested in such a system, those which bind $\beta\gamma$ more effectively will sequester $\beta\gamma$ and reduce or block signaling. Preferably, such chimeras and *gpa1* subunits can be assessed in a *gpa1*⁻ background to avoid competition with *gpa1* for $\beta\gamma$. For example, G α s chimeric mutants (see below) carrying D229S, E10K, N254D, or S286P were found to sequester $\beta\gamma$ more effectively than the chimera with wild type sequences. Also, double mutants were even more effective than either single

5 mutant. Similarly, overexpression of $G_{\alpha s}$ by driving transcription from the highly efficient PGK promoter resulted in dampening of the receptor coupling which may be offset by introduction of the double mutant $G_{\alpha s}$ (D229S,E10K).

Guidance for making mutations in G protein subunits and in the construction of chimeric G protein subunits is provided below.

10

Site-Directed Mutagenesis versus Random Mutagenesis

There are numerous art recognized ways to solve the structure-function problems of the sort presented by attempts to define the determinants involved in mediating the association of the subunits that comprise the G protein heterotrimer. For example, in one approach, discussed above with respect to hybrid constructs, specific mutations or alterations are introduced into a molecule based upon the available experimental evidence. In a second approach, random mutagenesis techniques, coupled with selection or screening systems, are used to introduce large numbers of mutations into a molecule, and that collection of randomly mutated molecules is then subjected to a selection for the desired phenotype or a screen in which the desired phenotype can be observed against a background of undesirable phenotypes.

With random mutagenesis one can mutagenize an entire molecule or one can proceed by cassette mutagenesis. In the former instance, the entire coding region of a molecule is mutagenized by one of several methods (chemical, PCR, doped oligonucleotide synthesis) and that collection of randomly mutated molecules is subjected to selection or screening procedures. Random mutagenesis can be applied in this way in cases where the molecule being studied is relatively small and there are powerful and stringent selections or screens available to discriminate between the different classes of mutant phenotypes that will inevitably arise. In the second approach, discrete regions of a protein, corresponding either to defined structural (i.e. α -helices, β -sheets, turns, surface loops) or functional determinants (e.g., catalytic clefts, binding determinants, transmembrane segments) are subjected to saturating or semi-random mutagenesis and these mutagenized cassettes are re-introduced into the context of the otherwise wild type allele.

35 Cassette mutagenesis is most useful when there is experimental evidence available to suggest a particular function for a region of a molecule and there is a powerful selection and/or screening approach available to discriminate between interesting and uninteresting mutants. Cassette mutagenesis is also useful when the parent molecule is comparatively large and the desire is to map the functional domains

5 of a molecule by mutagenizing the molecule in a step-wise fashion, i.e. mutating one linear cassette of residues at a time and then assaying for function.

The present invention provides for applying random mutagenesis in order to further delineate the determinants involved in $G\alpha$ - $G\beta\gamma$ or subunit-receptor association. Random mutagenesis may be accomplished by many means, including:

10 1. PCR mutagenesis, in which the error prone Taq polymerase is exploited to generate mutant alleles of G protein subunits, which are assayed directly in yeast for an ability to couple.

2. Chemical mutagenesis, in which expression cassettes encoding G protein subunits are exposed to mutagens and the protein products of the mutant sequences are
15 assayed directly in yeast for an ability to couple.

3. Doped synthesis of oligonucleotides encoding portions of the G protein subunit gene.

4. *In vivo* mutagenesis, in which random mutations are introduced into the coding region of G protein subunits by passage through a mutator strain of *E. coli*, XL1-
20 Red (mutD5 mutS mutT) (Stratagene, Menasa, WI).

In certain embodiments, for example, the random mutagenesis may be focused on regions suspected to be involved in $G\alpha$ - $G\beta\gamma$ association. Random mutagenesis approaches are feasible for two reasons. First, in yeast one has the ability to construct stringent screens and facile selections (growth vs. death, transcription vs. lack of
25 transcription) that are not readily available in mammalian systems. Second, when using yeast it is possible to screen efficiently through thousands of transformants rapidly. For example, this relatively small region of $G\alpha$ subunits represents a reasonable target for cassette mutagenesis. Another region that may be amenable to cassette mutagenesis is that defining the surface of the switch region of $G\alpha$ subunits that is solvent-exposed in
30 the crystal structures of $G\alpha_i$ and transducin. From the data described below, this surface may contain residues that are in direct contact with yeast $G\beta\gamma$ subunits, and may therefore be a reasonable target for mutagenesis.

A. Modification of $G\alpha$

35 Some aspects of $G\alpha$ structure are relevant to the design of modified $G\alpha$ subunits. Alignments of $G\alpha$ and GPA1 can be made to determine sequence similarity. For alignments of the entire coding regions of GPA1 with $G\alpha_s$, $G\alpha_i$, and $G\alpha_O$, $G\alpha_q$ and $G\alpha_z$, see Dietzel and Kurjan (1987, *Cell* 50:573) and Lambright, et al. (1994, *Nature* 369:621-628). Additional sequence information is provided by Mattera, et al. (1986,
40 *FEBS Lett* 206:36-41), Bray, et al. (1986, *Proc. Natl. Acad. Sci USA* 83:8893-8897) and

5 Bray, et al. (1987, *Proc Natl. Acad Sci USA* 84:5115-5119). An alignment of GPA1 and
four other G α proteins is provided by Stone and Reed (1990. *Mol. Cell Biol.* 10:4439).

The gene encoding a G protein homolog of *S. cerevisiae* was cloned independently by Dietzel and Kurjan (*supra*) (who referred to the gene as SCG1) and by Nakafuku, et al. (1987 *Proc Natl Acad Sci* 84:2140-2144) (who called the gene GPA1). Sequence analysis revealed a high degree of homology between the protein encoded by this gene and mammalian G α . GPA1 encodes a protein of 472 amino acids, as compared with approximately 340-350 amino acids for most mammalian G α subunits in four described families, G α s, G α i, G α q and G α 12/13. Nevertheless, GPA1 shares overall sequence and structural homology with all G α proteins identified to date. The highest overall homology in GPA1 is to the G α i family (48% identity, or 65% with conservative substitutions) and the lowest is to G α s (33% identity, or 51% with conservative substitutions) (Nakafuku, et al., *supra*).

The regions of high sequence homology among G α subunits are dispersed throughout their primary sequences, with the regions sharing the highest degree of homology mapping to sequence that comprises the guanine nucleotide binding/GTPase domain. This domain is structurally similar to the $\alpha\beta$ fold of ras proteins and the protein synthesis elongation factor EF-Tu. This highly conserved guanine nucleotide-binding domain consists of a six-stranded β sheet surrounded by a set of five α -helices. It is within these β sheets and α helices that the highest degree of conservation is observed among all G α proteins, including GPAI. The least sequence and structural homology is found in the intervening loops between the β sheets and α helices that define the core GTPase domain. There are a total of four "intervening loops" or "inserts" present in all G α subunits. In the crystal structures reported to date for the GDP- and GTP γ S-liganded forms of bovine rod transducin (Noel, et al. (1993) *Nature* 366:654-663); (Lambright, et al. (1994) *Nature* 369:621-628), the loop residues are found to be outside the core GTPase structure. Functional roles for these loop structures have been established in only a few instances. A direct role in coupling to phosphodiesterase- γ has been demonstrated for residues within inserts 3 and 4 of Gat (Rarick, et al. (1992) *Science* 256:1031-1033); (Artemyev, et al. (1992) *J. Biol. Chem.* 267:25067-25072), while a "GAP-like" activity has been ascribed to the largely α -helical insert 1 domain of G α S (Markby, et al. (1993) *Science* 262:1805-1901).

While the amino- and carboxy-termini of G α subunits do not share striking homology either at the primary, secondary, or tertiary levels, there are several generalizations that can be made about them. First, the amino termini of G α subunits have been implicated in the association of G α with G $\beta\gamma$ complexes and in membrane association via N-terminal myristoylation. In addition, the carboxy-termini have been

5 implicated in the association of $G\alpha\beta\gamma$ heterotrimeric complexes with G protein-coupled receptors (Sullivan, et al. (1987) *Nature* 330:758-760); West, et al. (1985) *J. Biol. Chem.* 260:14428-14430); (Conklin, et al. (1993) *Nature* 363:274-276); (Kallal and Kurjan. 1997. *Mol. Cell. Biol.* 17:2897). Data in support of these generalizations about the
 10 function of the N-terminus derive from several sources, including both biochemical and genetic studies.

In the GPA4I $G\alpha$ hybrids, the amino terminal 41 residues are derived from GPA1. All residues following position 41 are contributed by the human $G\alpha$ subunits, including the consensus nucleotide binding motif. For alignments of the entire coding regions of GPA1 with $G\alpha_s$, $G\alpha_i$, and $G\alpha_o$, $G\alpha_q$ and $G\alpha_z$, see Dietzel and Kurjan (1987, *Cell*
 15 50:573) and Lambright, et al. (1994, *Nature* 369:621-628). Additional sequence information is provided by Mattera, et al. (1986, *FEBS Lett* 206:36-41), Bray, et al. (1986, *Proc. Natl. Acad. Sci USA* 83:8893-8897) and Bray, et al. (1987, *Proc Natl. Acad Sci USA* 84:5115-5119).

There is little if any sequence homology shared among the amino termini of $G\alpha$
 20 subunits. The amino terminal domains of $G\alpha$ subunits that precede the first β -sheet vary in length from 41 amino acids (GPA1) to 31 amino acids ($G\alpha_t$). Most $G\alpha$ subunits share the consensus sequence for the addition of myristic acid at their amino termini, although not all $G\alpha$ subunits that contain this motif have myristic acid covalently associated with the glycine at position 2 (Speigel, et al. (1991) *TIBS* 16:338-3441). The role of this post-
 25 translational modification has been inferred from studies in which the activity of mutant $G\alpha$ subunits from which the consensus sequence for myristoylation has been added or deleted has been assayed (Mumby et al. (1990) *Proc. Natl. Acad. Sci. USA* 87: 728-732; (Linder, et al. (1991) *J. Biol Chem.* 266:4654-4659); Gallego, et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:9695-9699). These studies suggest two roles for N-terminal
 30 myristoylation. First, the presence of amino-terminal myristic acid has in some cases been shown to be required for association of $G\alpha$ subunits with the membrane, and second, this modification has been demonstrated to play a role in modulating the association of $G\alpha$ subunits with $G\beta\gamma$ complexes. The role of myristoylation of the GPA1 gene products, at present is unknown.

35 In other biochemical studies aimed at examining the role of the amino-terminus of $G\alpha$ in driving the association between $G\alpha$ and $G\beta\gamma$ subunits, proteolytically or genetically truncated versions of $G\alpha$ subunits were assayed for their ability to associate with $G\beta\gamma$ complexes, bind guanine nucleotides and/or to activate effector molecules. In all cases, $G\alpha$ subunits with truncated amino termini were deficient in all three functions
 40 (Graf, et al. (1992) *J. Biol. Chem.* 267:24307-24314); (Journot, et al. (1990) *J. Biol. Chem.* 265:9009-9015); and (Neer, et al. (1988) *J. Biol. Chem* 263:8996-9000). Slepak,

et al. (1993, *J. Biol. Chem.* 268:1414-1423) reported a mutational analysis of the N-terminal 56 a.a. of mammalian G α_o expressed in *Escherichia coli*. Molecules with an apparent reduced ability to interact with exogenously added mammalian G $\beta\gamma$ were identified in the mutant library. As the authors pointed out, however, the assay used to screen the mutants the extent of ADP-ribosylation of the mutant G α by pertussis toxin was not a completely satisfactory probe of interactions between G α and G $\beta\gamma$. Mutations identified as inhibiting the interaction of the subunits, using this assay, may still permit the complexing of G α and G $\beta\gamma$ while sterically hindering the ribosylation of G α by toxin. Other work has revealed specific amino acid residues of GPA1 that are important in GPA1 function. For example, a E307K mutation appears to create an α subunit with a broadened specificity for G β subunits (Whiteway et al. 1994. *Mol. Cell. Biol.* 14:3223). Interestingly, the residue in the mammalian G α subunit which is equivalent to the E307 position is diagnostic for a particular class of mammalian α subunits. For example, the G $_s\alpha$ subunits contain a lysine at this position, the G $_o$ and G $_i$ α subunits contain a histidine, the transducin α subunits have a glutimine, the Gq α subunits have a proline, and the G $_{13}$ α subunits have an aspartic acid at this site (Whiteway et al. *supra*).

Genetic studies examined the role of amino-terminal determinants of G α in heterotrimer subunit association have been carried out in both yeast systems using GPA1-mammalian G α hybrids (Kang, et al. (1990) *Mol. Cell. Biol.* 10:2582-2590) and in mammalian systems using G α_i /G α_s hybrids (Russell and Johnson (1993) *Mol. Pharmacol.* 44:255-263). In the former studies, gene fusions, composed of yeast GPA1 and mammalian G α sequences were constructed by Kang, et al. (*supra*) and assayed for their ability to complement a *gpa1* null phenotype (i.e., constitutive activation of the pheromone response pathway) in *S. cerevisiae*. Kang, et al. demonstrated that wild type mammalian G α_s , G α_i but not G α_o proteins are competent to associate with yeast G α and suppress the *gpa1* null phenotype, but only when overexpressed. Fusion proteins containing the amino-terminal 330 residues of GPA1 sequence linked to 160, 143, or 142 residues of the mammalian G α_s , G α_i and G α_o carboxyl-terminal regions, respectively, also coupled to the yeast mating response pathway when overexpressed on high copy plasmids with strong inducible (CUP) or constitutive (PGK) promoters. All three of these hybrid molecules were able to complement the *gpa1* null mutation in a growth arrest assay, and were additionally able to inhibit α -factor responsiveness and mating in tester strains. These last two observations argue that hybrid yeast-mammalian G α subunits are capable of interacting directly with yeast G $\beta\gamma$, thereby disrupting the normal function of the yeast heterotrimer. Fusions containing the amino terminal domain of G α_s , G α_i or G α_o , however, did not complement the *gpa1* null phenotype, indicating a requirement for determinants in the amino terminal 330 amino acid residues of GPA1 for association and sequestration of yeast G $\beta\gamma$ complexes. Taken together,

5 these data suggest that determinants in the amino terminal region of $G\alpha$ subunits determine not only the ability to associate with $G\beta\gamma$ subunits in general, but also with specific $G\beta\gamma$ subunits in a species-restricted manner.

Hybrid $G\alpha i/G\alpha s$ subunits have been assayed in mammalian expression systems (Russell and Johnson (*supra*). In these studies, a large number of chimeric $G\alpha$ subunits
 10 were assayed for an ability to activate adenylyl cyclase, and therefore, indirectly, for an ability to interact with $G\beta\gamma$ (i.e., coupling of $G\alpha$ to $G\beta\gamma$ = inactive cyclase; uncoupling of $G\alpha$ from $G\beta\gamma$ = active cyclase). From these studies a complex picture emerged in which determinants in the region between residues 25 and 96 of the hybrids were found to determine the state of activation of these alleles as reflected in their rates of guanine
 15 nucleotide exchange and GTP hydrolysis and the extent to which they activated adenylyl cyclase *in vivo*. These data could be interpreted to support the hypothesis that structural elements in the region between the amino terminal methionine and the β sheet identified in the crystal structure of $G\alpha t$ (see Noel, et al. *supra* and Lambright, et al. *supra*) are involved in determining the state of activity of the heterotrimer by (1) driving
 20 association/dissociation between $G\alpha$ and $G\beta\gamma$ subunits; (2) driving GDP/GTP exchange. While there is no direct evidence provided by these studies to support the idea that residues in this region of $G\alpha$ and residues in $G\beta\gamma$ subunits contact one another, the data nonetheless provide a positive indication for the construction of hybrid $G\alpha$ subunits that retain function. There is, however, a negative indicator that derives from this work in
 25 that some hybrid constructs resulted in constitutive activation of the chimeric proteins (i.e., a loss of receptor-dependent stimulation of $G\beta\gamma$ dissociation and effector activation).

B. Construction of chimeric $G\alpha$ subunits.

30 In preferred embodiments chimeric $G\alpha$ subunits retain as much of the sequence of the native mammalian proteins as possible and, in particularly preferred embodiments, the level of expression for the heterologous components should approach, as closely as possible, the level of their endogenous counterparts. The results described by King, et al. (1990, Science 250:121-123) for expression of the human $\beta 2$ -adrenergic
 35 receptor and $G\alpha s$ in yeast, taken together with negative results obtained by Kang, et al. (*supra*) with full-length mammalian $G\alpha$ subunits other than $G\alpha s$, led to the following preferred embodiments for the development of yeast strains in which mammalian G protein-coupled receptors could be linked to the pheromone response pathway.

40 In one embodiment, mammalian $G\alpha$ subunits are expressed using the native sequence of each subunit or, alternatively, as minimal gene fusions with sequences from the amino-terminus of GPA1 replacing the homologous residues from the mammalian

G α subunits. In another embodiment, mammalian G α subunits are expressed from the GPA1 promoter either on low copy plasmids or after integration into the yeast genome as a single copy gene. In certain embodiments, endogenous G $\beta\gamma$ subunits are provided by the yeast STE4 and STE18 loci, while in other embodiments chimeric or heterologous G β and/or G γ subunits are also provided.

C. Rational Design of Chimeric G α Subunits

Several classes of rationally designed GPA1-mammalian G α hybrid subunits have been tested for the ability to couple to yeast $\beta\gamma$. The first, and largest, class of hybrids are those that encode different lengths of the GPA1 amino terminal domain in place of the homologous regions of the mammalian G α subunits. This class of hybrid molecules includes GPABAMH1, GPA41, GPAID, and GPALW hybrids, described below. The rationale for constructing these hybrid G α proteins is based on results, described above, that bear on the importance of the amino terminal residues of G α in mediating interaction with G $\beta\gamma$.

Preferably, the yeast G α subunit is replaced by a chimeric G α subunit in which a portion, e.g., at least about 20, more preferably at least about 40, amino acids, from the amino terminus of the yeast G α , is fused to a sequence from a mammalian (or other exogenous) G α . While about 40 amino acids is the suggested starting point, shorter or longer portions may be tested to determine the minimum length required for coupling to yeast G $\beta\gamma$ and the maximum length compatible with retention of coupling to the exogenous receptor. It is presently believed that only the final 10 or 20 amino acids at the carboxy terminus of the G α subunit are required for interaction with the receptor.

i. GPABAMH1 hybrids.

Kang et al. *supra*. described hybrid G α subunits encoding the amino terminal 310 residues of GPA1 fused to the carboxyl terminal 160, 143 and 142 residues, respectively, of G α S, G α i2, and G α o. In all cases examined by Kang et al., the hybrid proteins were able to complement the growth arrest phenotype of *gpa1* strains. Hybrids between GPA1 and G α i3, G α q and G α 16 can be constructed, as described below, and functionally complement the growth arrest phenotype of *gpa1* strains.

GPA41 hybrids: The rationale for constructing a minimal hybrid encoding only 41 amino acids of GPA1 relies upon the biochemical evidence for the role of the amino-terminus of G α subunits discussed above, together with the following observation. G β and G γ subunits are known to interact via α -helical domains at their respective amino-termini (Pronin, et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6220-6224);

5 Garritsen, et al.1993). The suggestion that the amino termini of G α subunits may form an helical coil and that this helical coil may be involved in association of G α with G $\beta\gamma$ (Masters et al (1986) *Protein Engineering* 1:47-54); Lupas et al.(1992) *FEBS Lett.* 314:105-108) leads to the hypothesis that the three subunits of the G-protein heterotrimer interact with one another reversibly through the winding and unwinding of
 10 their amino-terminal helical regions. (For further description of the crystal structure of G proteins, and subunits thereof, see Lambright et al. (1996) *Nature* 379:311-319 and Sondek et al. (1996) *Nature* 379:369-374.) A mechanism of this type has been suggested, as well, from an analysis of leucine zipper mutants of the GCN4 transcription factor (Harbury, et al. (1993) *Science* 262:1401-1407). The rationale for constructing
 15 hybrids like those described by Kang, et al. *supra.*, that contain a majority of yeast sequence and only minimal mammalian sequence, derives from their ability to function in assays of coupling between G α and G $\beta\gamma$ subunits. However, these chimeras had never been assayed for an ability to couple to both mammalian G protein-coupled receptors and yeast G $\beta\gamma$ subunits, and hence to reconstitute a hybrid signaling pathway
 20 in yeast.

GPA₄₁ hybrids that have been constructed and tested include G α_s , G α_i2 , G α_i3 , G α_q , G α_{o_a} , G α_{o_b} and G α_{16} . Hybrids of G α_s , G α_i2 , G α_i3 , and G α_{16} functionally complement the growth arrest phenotype of *gal* strains, while GPA₄₁ hybrids of G α_{o_a} and G α_{o_b} do not. In addition to being tested in a growth arrest assay, these constructs
 25 have been assayed in the more sensitive transcriptional assay for activation of a *fusl*-HIS3 gene. In both of these assays, the GPA₄₁-G α_s hybrid couples less well than the GPA₄₁-i2, -i3, and -16 hybrids, while the GPA₄₁ -o_a, and -o_b hybrids do not function in either assay.

Several predictive algorithms indicate that the amino terminal domain up to the
 30 highly conserved sequence motif LLLLGAGESG (SEQ ID NO: 12) (the first L in this motif is residue 43 in GPA1) forms a helical structure with amphipathic character. Assuming that a heptahelical repeat unit, the following hybrids between yeast GPA1 and mammalian G α_s can be used to define the number of helical repeats in this motif necessary for hybrid function:

35 GPA1-7/G α_s 8-394
 GPA1-14/G α_s 15-394
 GPA1-21/G α_s 22-394
 GPA1-28/G α_s 29-394
 GPA1-35/G α_s 36-394
 40 GPA1-42/G α_s 43-394

5 In these hybrids, the prediction is that the structural repeat unit in the amino terminal domain up to the tetra-leucine motif is 7, and that swapping sequences in units of 7 will in effect amount to a swap of unit turns of turns of the helical structure that comprises this domain.

10 A second group of "double crossover" hybrids of this class are those that are aligned on the first putative heptad repeat beginning with residue G11 in GPA1. In these hybrids, helical repeats are swapped from GPA1 into a GaS backbone one heptad repeat unit at a time.

GαS1-10/GPA11-17/GαS18-394

GαS1-17/GPA18-24/GαS25-394

15 GαS1-17/GPA25-31/GαS32-394

GαS1-17/GPA32-38/GαS39-394

20 The gap that is introduced between residues 9 and 10 in the GαS sequence is to preserve the alignment of the LLLLGAGE (SEQ ID NO: 13) sequence motif. This class of hybrids can be complemented by cassette mutagenesis of each heptad repeat followed by screening of these collections of "heptad" libraries in standard coupling assays.

25 A third class of hybrids based on the prediction that the amino terminus forms a helical domain with a heptahelical repeat unit are those that effect the overall hydrophobic or hydrophilic character of the opposing sides of the predicted helical structure (See Lupas et al. *supra*). In this model, the a and d positions of the heptad repeat abcdefg are found to be conserved hydrophobic residues that define one face of the helix, while the e and g positions define the charged face of the helix. In this class of hybrids, the sequence of the GαS parent is maintained except for specific substitutions at one or more of the following critical residues to render the different helical faces of GαS more "GPA1-like"

30

K8Q

+I-10

E10G

Q12E

35

R13S

N14D

E15P

E15F

K17L

40

E21R

5 K28Q
K32L
V36R

This collection of single mutations could be screened for coupling efficiency to yeast G $\beta\gamma$ and then constructed in combinations (double and greater if necessary).

10 A fourth class of hybrid molecules that span this region of GPA1-G α hybrids are those that have junctions between GPA1 and G α subunits introduced by three primer PCR. In this approach, the two outside primers are encoded by sequences at the initiator methionine of GPA1 on the 5' side and at the tetra-leucine motif of G α S (for example) on the 3' side. A series of junctional primers spanning different junctional points can be
15 mixed with the outside primers to make a series of molecules each with different amounts of GPA1 and G α S sequences, respectively.

ii. GPA_{ID} and GPA_{LW} hybrids.

The regions of high homology among G $\beta\gamma$ subunits that have been identified by
20 sequence alignment are interspersed throughout the molecule. The G1 region containing the highly conserved GSGESGDST (SEQ ID NO: 14) motif is followed immediately by a region of very low sequence conservation, the "il" or insert 1 region. Both sequence and length vary considerably among the il regions of the G α subunits. By aligning the sequences of G α subunits, the conserved regions bounding the il region were identified
25 and two additional classes of GPA1-G α hybrids were constructed. The GPA_{ID} hybrids encode the amino terminal 102 residues of GPA1 (up to the sequence QARKLGIQ) (SEQ ID NO: 15) fused in frame to mammalian G α subunits, while the GPA_{LW} hybrids encode the amino terminal 244 residues of GPA1 (up to the sequence LIHEDIKA (SEQ ID NO: 16) in GPA1). The reason for constructing the GPA_{ID} and GPA_{LW}
30 hybrids was to test the hypothesis that the il region of GPA1 is required for mediating the interaction of GPA1 with yeast G $\beta\gamma$ subunits, for the stable expression of the hybrid molecules, or for function of the hybrid molecules. The GPA_{ID} hybrids contain the amino terminal domain of GPA1 fused to the il domain of mammalian subunits, and therefore do not contain the GPA1 il region, while the GPA_{LW} hybrids contain the
35 amino terminal 244 residues of GPA1 including the entire il region (as defined by sequence alignments). Hybrids of both GPA_{ID} and GPA_{LW} classes were constructed for G α S, C- α i2, G α i3, G α o_a, and G α 16; none of these hybrids complemented the gpa1 growth arrest phenotype.

Subsequent to the construction and testing of the GPA_{ID} and GPA_{LW} classes of
40 hybrids, the crystal structures of G_{transducin} in both the GDP and GTP γ S-liganded

5 form, and the crystal structure of several G α il variants in the GTP γ S-liganded and GDP-AlF₄ forms were reported (Noel et al. *supra*; Lambright et al. *supra*; and Coleman et al.(1994) *Science* 265:1405-1412). The crystal structures reveal that the il region defined by sequence alignment has a conserved structure that is comprised of six alpha helices in a rigid array, and that the junctions chosen for the construction of the GPA_{ID} and GPA_{LW} hybrids were not compatible with conservation of the structural features of
10 the il region observed in the crystals. The junction chosen for the GPA_{ID} hybrids falls in the center of the long α A helix; chimerization of this helix in all likelihood destabilizes it and the protein structure in general. The same is true of the junction chosen for the GPA_{LW} hybrids in which the crossover point between GPA1 and the
15 mammalian G α subunit falls at the end of the short α C helix and therefore may distort it and destabilize the protein.

The failure of the GPA_{ID} and GPA_{LW} hybrids is predicted to be due to disruption of critical structural elements in the il region as discussed above. Based upon new alignments and the data presented in Noel et al (*supra*), Lambright et al (*supra*), and
20 Coleman et al (*supra*), this problem can be averted with the ras-like core domain and the il helical domain are introduced outside of known structural elements like alpha-helices.

Hybrid A G α S1-67/GPA66-299/G α S203-394

This hybrid contains the entire il insert of GPA1 interposed into the G α S sequence.

25

Hybrid B GPA1-41/G α S4443-67/GPA66-299/G α S203-394

This hybrid contains the amino terminal 41 residues of GPA1 in place of the 42 amino terminal residues of G α S found in Hybrid A.

30 *iii. G α s Hybrids.*

There is evidence that the "switch region" encoded by residues 171-237 of G α transducin (using the numbering of (Noel et al (*supra*)) also plays a role in G $\beta\gamma$ coupling. First, the G226A mutation in G α S prevents the GTP-induced conformational change that occurs with exchange of GDP for GTP upon receptor activation by ligand.
35 This residue maps to the highly conserved sequence DVGGQ (SEQ ID NO: 17), present in all G α subunits and is involved in GTP hydrolysis. In both the G α t and G α il crystal structures, this sequence motif resides in the loop that connects the β 3 sheet and the α 2 helix in the guanine nucleotide binding core. In addition to blocking the conformational change that occurs upon GTP binding, this mutation also prevents dissociation of GTP-liganded G α s from G $\beta\gamma$.
40 Second, crosslinking data reveals that a highly conserved

5 cysteine residue in the $\alpha 2$ helix (C215 in $G\alpha o$, C210 in $G\alpha t$) can be crosslinked to the carboxy terminal region of $G\beta$ subunits. Finally, genetic evidence (Whiteway et al. (1993) Mol Cell Biol. 14:3233-3239) identifies an important single residue in GPA1 (E307) in the $\beta 2$ sheet of the core structure that may be in direct contact with $\beta \gamma$. A mutation in the GPA1 protein at this position suppresses the constitutive signaling
 10 phenotype of a variety of STE4 ($G\beta$) dominant negative mutations that are also known to be defective in $G\alpha$ - $G\beta \gamma$ association (as assessed in two-hybrid assay in yeast as well as by more conventional genetic tests).

The GPA1 switch region suppresses coupling to yeast $G\beta \gamma$ (SGS), while in the context of the GPA1 amino terminus the GPA1 switch region stabilizes coupling with
 15 $G\beta \gamma$ ($GP\beta \gamma$ -SGS). This suggests that these two regions of GPA1 collaborate to allow interactions between $G\alpha$ subunits and $G\beta \gamma$ subunits. This conclusion is somewhat mitigated by the observation that the GPA₄₁- $G\alpha s$ hybrid that does not contain the GPA1 switch region is able to complement the growth arrest phenotype of *gal* strains.

The role of the surface-exposed residues of this region may be crucial for
 20 effective coupling to yeast $G\beta \gamma$, and can be incorporated into hybrid molecules as follows below.

$G\alpha S$ -GPA-Switch $G\alpha S$ 1-202/GPA298-350/ $G\alpha S$ 253-394

This hybrid encodes the entire switch region of GPA 1 in the context of $G\alpha S$.

25

$G\alpha S$ -GPA- $\alpha 2$ GQS 1-226/GPA322-332/GQS 238-394

This hybrid encodes the $\alpha 2$ helix of GPA1 in the context of $G\alpha S$.

GPA41- $G\alpha S$ -GPA- $\alpha 2$ GPA1-41/GQS43-226/GPA322-332/GQS238-394

30 This hybrid encodes the 41 residue amino terminal domain of GPA1 and the $\alpha 2$ helix of GPA1 in the context of $G\alpha S$.

In addition, hybrids that alter the surface exposed residues of the $\beta 2$ and $\beta 3$ sheets of αS so that they resemble those of the GPA1 QS helix can be made. These
 35 altered $\alpha 2$ helical domains have the following structure. (The positions of the altered residues correspond to $G\alpha S$.)

L203K

K211E

D215G
K216S
D229S

These single mutations can be engineered into a G α S backbone singly and in pairwise combinations. In addition, they can be introduced in the context of both the full length G α S and the GPA₄₁-G α S hybrid described previously. All are predicted to improve the coupling of G α subunits to yeast G $\beta\gamma$ subunits by virtue of improved electrostatic and hydrophobic contacts between this region and the regions of G β defined by Whiteway and co-workers (Whiteway et al (*supra*) that define site(s) that interact with GPA1).

In summary, the identification of hybrid G α subunits that couple to the yeast pheromone pathway has led to the following general observations. First, GPABAMH1 hybrids associate with yeast G $\beta\gamma$, therefore at a minimum these hybrids contain the determinants in GPA1 necessary for coupling to the pheromone response pathway. Second, the amino terminal 41 residues of GPA1 contain sufficient determinants to facilitate coupling of G α hybrids to yeast G $\beta\gamma$ in some, but not all, instances, and that some G α subunits contain regions outside of the first 41 residues that are sufficiently similar to those in GPA1 to facilitate interaction with GPA1 even in the absence of the amino terminal 41 residues of GPA1. Third, there are other determinants in the first 310 residues of GPA1 that are involved in coupling G α subunits to yeast G $\beta\gamma$ subunits.

The various classes of hybrids noted above are not mutually exclusive. For example, a GPA1 containing GPA1-41 could also feature the L203K mutation.

While, for the sake of simplicity, hybrids of yeast GPA1 and a mammalian G α s have been described here, it will be appreciated that hybrids may be made of other yeast G α subunits and/or other mammalian G α subunits, notably mammalian G α i subunits. Moreover, while the described hybrids are constructed from two parental proteins, hybrids of three or more parental proteins are also possible.

As shown in the Examples, chimeric G α subunits have been especially useful in coupling receptors to G α i species.

iv. Expression of G α

Kang et al. *supra* reported that several classes of native mammalian G α subunits were able to interact functionally with yeast α subunits when expression of G α was driven from a constitutively active, strong promoter (PGK) or from a strong inducible promoter (CUP). These authors reported that rat G α S, G α i2 or G α o expressed at high level coupled to yeast $\beta\gamma$. High level expression of mammalian G α (i.e. non-

5 stoichiometric with respect to yeast $\beta\gamma$) is not preferred for uses like those described in this application. Reconstruction of G protein-coupled receptor signal transduction in yeast requires the signaling component of the heterotrimeric complex ($G\beta\gamma$) to be present stoichiometrically with $G\alpha$ subunits. An excess of $G\alpha$ subunits (as was required for coupling of mammalian $G\alpha i2$ and $G\alpha o$ to yeast $G\beta\gamma$ in Kang et al.) would dampen
 10 the signal in systems where $G\beta\gamma$ subunits transduce the signal. An excess of $G\alpha$ subunits raises the background level of signaling in the system. Preferably, levels of $G\alpha$ and $G\beta\gamma$ subunits are balanced. For example, heterologous $G\alpha$ subunits may be expressed from a low copy (CEN ARS) vector containing the endogenous yeast GPA1 promoter and the GPA1 3' untranslated region. The minimum criterion, applied to a
 15 heterologous $G\alpha$ subunit with respect to its ability to couple functionally to the yeast pheromone pathway, is that it complement a *gpa1* genotype when expressed from the GPA1 promoter on low copy plasmids or from an integrated, single copy gene. In the work described in this application, heterologous $G\alpha$ subunits have been assayed in two biological systems. In the first assay heterologous $G\alpha$ subunits are tested for an ability
 20 to functionally complement the growth arrest phenotype of *gpa1* strains. In the second assay the transcription of a *fus1*-HIS3 reporter gene is used to measure the extent to which the pheromone response pathway is activated, and hence the extent to which the heterologous $G\alpha$ subunit sequesters the endogenous yeast $G\beta\gamma$ complex. Mammalian $G\alpha s$, $G\alpha i2$, $G\alpha i3$, $G\alpha q$, $G\alpha 11$, $G\alpha 16$, $G\alpha o_a$, $G\alpha o_b$, and $G\alpha z$ from rat, murine or human
 25 origins were expressed from a low copy, CEN ARS vector containing the GPA1 promoter. Functional complementation of *gpa1* strains was not observed in either assay system with any of these full-length $G\alpha$ constructs with the exception of rat and human $G\alpha S$.

30 *D. Chimeric Yeast $\beta\gamma$ subunits*

In addition to or in place of modifying G protein $G\alpha$ subunits, yeast or heterologous $G\beta$ or $G\gamma$ subunits can be modified. The methods described above with regard to $G\alpha$ modification can be used to alter either or both of these subunits as well. For example, alignments of the yeast sequence and heterologous sequences can be made
 35 and combined with information regarding important functional domains. Such information can then be used to provide guidance in making mutations in yeast or heterologous sequences. Likewise, chimeric $G\beta$ or $G\gamma$ molecules can be constructed to enhance the coupling of heterologous GPCRs to a yeast pheromone signaling pathway.

The yeast STE4 and STE18 are related to the metazoan G protein β and γ
 40 subunits, respectively (Whiteway et al. 1989. Cell. 56:467). The β and γ subunits must be capable of interaction with one another as well as with the α subunit and with the

5 effector. Previous work has suggested that mammalian β or γ subunits are divergent enough from their yeast homologues that they cannot functionally replace STE4 or STE 18. (Coria et al. 1996. *Yeast*. 12:41). Thus, in preferred embodiments, modifications are made to heterologous G β or G γ subunits expressed in yeast and/or chimeric subunits are made to enhance heterologous receptor coupling.

10 The primary structure of G-protein β subunits is highly conserved from yeast to humans; Ste4 shares approximately 40% identity with human G β isoforms (Leberer et al. 1992 *EMBO Journal* 11:4085). STE 4 and the G β s are 420, and 340 or 341 amino acids long, respectively, and belong to the family of proteins with WD-40 motifs (van der Voorn and Ploegh. 1992. *FEBS Lett.* 307:131). These motifs can be used to divide
15 G β and STE4 into eight blocks (Coria et al. *Yeast* 1996. 12:41). Among the mammalian G β s, some have been found to exhibit G γ subunit selectivity (Pronin and Gautham. 1992. *Proc. Natl. Acad. Sci. USA* 89:6220; Schmidt et al. 1992. *J. Biol. Chem.* 267:13807; Kleuss et al. 1992. *Nature*. 358:424). An alignment of the metazoan and yeast G protein β subunits is provided by Corai et al. (1996. *Yeast*.
20 12:41). Such an alignment can be used to provide guidance for making mutations to G protein β subunits as described for G α above. In addition, certain regions of STE4 have been found to be important and thus, may be less amenable to manipulation than other portions of the polypeptide. For example, the c-terminus of the STE4 product is essential for downstream signaling (Coria et al. 1995. *FEBS Letters* 367:122).
25 Mutations to two small regions in the amino terminal half of Ste4 have also been shown to inhibit signaling (Leberer et al. *supra*). Mutations which influence the interaction of STE4 and GPA1 have also been identified; mutations to the second copy of the WD40 repeat can be modified to reduce the interaction between STE4 and GPA1, without influencing other aspects of STE4 function (Whiteway et al. 1994. *Mol. Cell. Biol.*
30 14:3223)

The G γ s, including STE18, diverge more strongly from each other than do the G β s. Even among the mammalian G protein γ subunits, there is a fair amount of divergence. The γ subunit may determine the functional specificity of the $\beta\gamma$ subunit complex. Complete cDNAs for the γ 1 subunit from bovine retina (Hurley et al. *Proc.*
35 *Nat'l Acad. Sci USA*. 1984. 81:6948) the γ 1, γ 3, and γ 7 subunits from bovine brain (Robishaw et al. *J. Biol. Chem.* 1989. 264:15758; Gautam et al. *Science*. 1989. 244:971; Gautam et al. *Proc. Nat'l Acad. Sci. USA*. 1990 87:7973; Cali et al. *J. Biol. Chem.* 1992. 267:24023), and the γ 5 subunit from bovine and rat liver (Gisher et al. 1992. 12:1585) have been reported.

40 The STE18 gene of yeast terminates with a CAAX box (SEQ ID NO: 18) (where A is an aliphatic amino acid, and X is any uncharged amino acid). This sequence is

5 involved in prenylation of G γ and is likely important in the localization of G γ to the membrane and may, thus, be less amenable to manipulation than other portions of the sequence. (Kurjan. 1992. *Ann. Rev. Biochem.* 61:1097). Saturation mutagenesis has also provided insight into regions of STE18 that are important in STE18 function. Mutations in STE18 which compensate for mutations in STE4 were identified at serine
 10 65, threonine 71, and valine 80. Dominant negative alleles of the STE18 gene were also identified (Whiteway et al. 1992. *Biochem. Cell. Biol.* 70:1230). These truncated proteins were found to lack the carboxyl terminus of STE 18, including the CAAX box (SEQ ID NO: 18) (Whiteway et al. *supra*).

An alignment of yeast G γ , STE18, and mammalian G γ s can be made as indicated
 15 for the other G protein subunits. Such an alignment can be used in constructing mutant G γ subunits or chimeric G γ subunits. In preferred embodiments, mammalian G γ 2 is used in making G protein γ subunit chimeras.

GPCR Expression

20 In other embodiments, a secretory signal of a yeast protein can be used to direct transport of the receptors, for example G protein coupled receptors, to the plasma membrane as described in the appended examples. Previous work has demonstrated the secretory expression of foreign proteins in yeast cells using the signal sequence of a yeast secreted protein such as invertase or acid phosphatase, encoded by the *SUC2* and
 25 *PHO5* genes, respectively (Schraber, M.D. et al. (1986) *Methods Enzymol.* 119:416; Moir, D.T. et al. (1991) *Methods Enzymol.* 194:491-507). The vast majority of the secreted proteins possess a hydrophobic N-terminal signal sequence which targets them to the endoplasmic reticulum. A leader sequence of the α -factor precursor encoded by the MF α 1 gene was shown to promote the most efficient secretion of various
 30 heterologous proteins. In addition to a signal sequence, the α -factor leader includes a hydrophilic pro-region which is believed to facilitate protein transport at the later stages of the secretory pathway.

Both secreted and membrane proteins including G protein coupled receptors are delivered to the cell surface through the same secretory pathway. Some receptors, for
 35 example, metabotropic glutamate receptors and vasoactive intestinal peptide receptors, also possess the N-terminal signal sequence, whereas some do not. In the latter case, a first transmembrane domain is believed to interact with the ER translocation machinery. The use of yeast secretory signals, in particular, the α -factor leader, may be desirable to provide the more efficient integration of the receptors into the membrane of the
 40 endoplasmic reticulum and transport to the plasma membrane. In fact, the cell surface

- 5 expression of the rat M5 receptor directed by the α -factor leader has been documented (Huang et al. (1992) *Biochem. Biophys. Res. Commun.* 181:1180).

VIII. Leader Sequences

It has been demonstrated that most of the mammalian extracellular, secreted proteins are poorly secreted when expressed in yeast. However, in many cases their secretion levels are markedly increased when their native signal sequences are replaced by the signal sequences of yeast proteins that interact more efficiently with the ER translocation complex. Specifically, the signal sequences of yeast invertase and acid phosphatase have been widely used in biotechnology to direct the secretory expression of the heterologous proteins. However, it is well established that even though many foreign proteins are targeted to the ER by the yeast signal sequences, not all of them advance further in the secretory pathway. The major problem consists in the misfolding and/or improper glycosylation of the heterologous proteins that results in their retention in the ER by the quality control apparatus of the yeast cell.

In many cases, a leader sequence of a precursor of yeast mating pheromone, α -factor, have been used successfully to overcome this problem (Brake, A.J. (1989) in *Yeast Genetic Engineering* (Barr, P.J., Brake, A.J., and Valenzuela, P., eds) pp. 269-280, Butterworths, London; Brake, A.J. (1990) *Meth. Enzymol.* 185, 408-441., and references cited therein). This sequence, in addition to the N-terminal signal peptide of 17 residues, includes a hydrophilic pro-region which contains 72 residues and bears three sites of N-linked glycosylation. The pro-region is extensively glycosylated in the ER and Golgi and is cleaved by Kex2 endopeptidase in the late Golgi compartment. The presence of the pro-region at the N-terminus is believed to allow some heterologous proteins to pass the quality control in the ER and to reach the periplasm. It is likely that the pro-region can somehow facilitate correct protein folding. Alternatively, it may be recognized by the quality control apparatus as a properly folded structural unit thus allowing an entire fusion protein to leave the ER.

The invertase leader can also be used. This leader sequence has been demonstrated to be cleaved from nascent invertase peptide, or nascent heterologous peptide, upon entrance into the endoplasmic reticulum, the apparent molecular weight of the receptor is consistent with this interpretation. The enzyme responsible for cleavage of the pre sequence, Kex2, resides in the trans Golgi.

5 In certain embodiments, such a leader sequence can be used to express a peptide library of the present invention. Yeast cells are bounded by a lipid bilayer called the plasma membrane. Between this plasma membrane and the cell wall is the periplasmic space. Peptides secreted by yeast cells cross the plasma membrane through a variety of mechanisms and thereby enter the periplasmic space. The secreted peptides are then free
10 to interact with other molecules that are present in the periplasm or displayed on the outer surface of the plasma membrane. The peptides then either undergo re-uptake into the cell, diffuse through the cell wall into the medium, or become degraded within the periplasmic space.

The test polypeptide library may be secreted into the periplasm by any of a
15 number of exemplary mechanisms, depending on the nature of the expression system to which they are linked. In one embodiment, the peptide may be structurally linked to a yeast signal sequence, such as that present in the α -factor precursor, which directs secretion through the endoplasmic reticulum and Golgi apparatus. Since this is the same route that the receptor protein follows in its journey to the plasma membrane,
20 opportunity exists in cells expressing both the receptor and the peptide library for a specific peptide to interact with the receptor during transit through the secretory pathway. This has been postulated to occur in mammalian cells exhibiting autocrine activation. Such interaction could yield activation of the response pathway during transit, which would still allow identification of those cells expressing a peptide agonist.
25 For situations in which peptide antagonists to externally applied receptor agonist are sought, this system would still be effective, since both the peptide antagonist and receptor would be delivered to the outside of the cell in concert. Thus, those cells producing an antagonist would be selectable, since the peptide antagonist would be properly and timely situated to prevent the receptor from being stimulated by the
30 externally applied agonist.

An alternative mechanism for delivering peptides to the periplasmic space is to use the ATP-dependent transporters of the STE6/MDR1 class. This transport pathway and the signals that direct a protein or peptide to this pathway are not as well characterized as is the endoplasmic reticulum-based secretory pathway. Nonetheless,
35 these transporters apparently can efficiently export certain peptides directly across the plasma membrane, without the peptides having to transit the ER/Golgi pathway. It is anticipated that at least a subset of peptides can be secreted through this pathway by expressing the library in context of the α -factor prosequence and terminal tetrapeptide. The possible advantage of this system is that the receptor and peptide do not come into
40 contact until both are delivered to the external surface of the cell. Thus, this system strictly mimics the situation of an agonist or antagonist that is normally delivered from

The present invention does not require periplasmic secretion of peptides, or, if such secretion is provided, any particular secretion signal or transport pathway. In certain embodiments, peptides expressed with a signal sequence may bind to and activate receptors prior to their transport to the cell surface.

The present invention does not require periplasmic secretion of peptides, or, if such secretion is provided, any particular secretion signal or transport pathway. In certain embodiments, peptides expressed with a signal sequence may bind to and

10 activate receptors prior to their transport to the cell surface.

λ_1°	λ_2°	λ_3°	λ_4°	λ_5°	λ_6°	λ_7°	λ_8°	λ_9°	λ_{10}°	λ_{11}°	λ_{12}°	λ_{13}°	λ_{14}°	λ_{15}°	λ_{16}°	λ_{17}°	λ_{18}°	λ_{19}°	λ_{20}°	λ_{21}°	λ_{22}°	λ_{23}°	λ_{24}°	λ_{25}°	λ_{26}°	λ_{27}°	λ_{28}°	λ_{29}°	λ_{30}°	λ_{31}°	λ_{32}°	λ_{33}°	λ_{34}°	λ_{35}°	λ_{36}°	λ_{37}°	λ_{38}°	λ_{39}°	λ_{40}°	λ_{41}°	λ_{42}°	λ_{43}°	λ_{44}°	λ_{45}°	λ_{46}°	λ_{47}°	λ_{48}°	λ_{49}°	λ_{50}°	λ_{51}°	λ_{52}°	λ_{53}°	λ_{54}°	λ_{55}°	λ_{56}°	λ_{57}°	λ_{58}°	λ_{59}°	λ_{60}°	λ_{61}°	λ_{62}°	λ_{63}°	λ_{64}°	λ_{65}°	λ_{66}°	λ_{67}°	λ_{68}°	λ_{69}°	λ_{70}°	λ_{71}°	λ_{72}°	λ_{73}°	λ_{74}°	λ_{75}°	λ_{76}°	λ_{77}°	λ_{78}°	λ_{79}°	λ_{80}°	λ_{81}°	λ_{82}°	λ_{83}°	λ_{84}°	λ_{85}°	λ_{86}°	λ_{87}°	λ_{88}°	λ_{89}°	λ_{90}°	λ_{91}°	λ_{92}°	λ_{93}°	λ_{94}°	λ_{95}°	λ_{96}°	λ_{97}°	λ_{98}°	λ_{99}°	λ_{100}°																																																																							
1.000000	0.999999	0.999998	0.999997	0.999996	0.999995	0.999994	0.999993	0.999992	0.999991	0.999990	0.999989	0.999988	0.999987	0.999986	0.999985	0.999984	0.999983	0.999982	0.999981	0.999980	0.999979	0.999978	0.999977	0.999976	0.999975	0.999974	0.999973	0.999972	0.999971	0.999970	0.999969	0.999968	0.999967	0.999966	0.999965	0.999964	0.999963	0.999962	0.999961	0.999960	0.999959	0.999958	0.999957	0.999956	0.999955	0.999954	0.999953	0.999952	0.999951	0.999950	0.999949	0.999948	0.999947	0.999946	0.999945	0.999944	0.999943	0.999942	0.999941	0.999940	0.999939	0.999938	0.999937	0.999936	0.999935	0.999934	0.999933	0.999932	0.999931	0.999930	0.999929	0.999928	0.999927	0.999926	0.999925	0.999924	0.999923	0.999922	0.999921	0.999920	0.999919	0.999918	0.999917	0.999916	0.999915	0.999914	0.999913	0.999912	0.999911	0.999910	0.999909	0.999908	0.999907	0.999906	0.999905	0.999904	0.999903	0.999902	0.999901	0.999900	0.999899	0.999898	0.999897	0.999896	0.999895	0.999894	0.999893	0.999892	0.999891	0.999890	0.999889	0.999888	0.999887	0.999886	0.999885	0.999884	0.999883	0.999882	0.999881	0.999880	0.999879	0.999878	0.999877	0.999876	0.999875	0.999874	0.999873	0.999872	0.999871	0.999870	0.999869	0.999868	0.999867	0.999866	0.999865	0.999864	0.999863	0.999862	0.999861	0.999860	0.999859	0.999858	0.999857	0.999856	0.999855	0.999854	0.999853	0.999852	0.999851	0.999850	0.999849	0.999848	0.999847	0.999846	0.999845	0.999844	0.999843	0.999842	0.999841	0.999840	0.999839	0.999838	0.999837	0.999836	0.999835	0.999834	0.999833	0.999832	0.999831	0.999

5 *B. GPCR Expression*

In other embodiments, a leader sequence of a yeast secreted protein can be used to direct transport of receptors, for example, G-protein coupled receptors to the plasma membrane as described in detail in the appended examples. Previous work has demonstrated the expression of foreign, secreted proteins in yeast cells using the α -factor leader. However, when a heterologous membrane bound receptor, the rat M5 receptor, was expressed using such a system, it was found that the heterologous GPCR did not functionally integrate into the yeast cell signaling pathway (Huang et al. *Biochem. and Biophys. Res. Comm.* 1992. 182:1180). The transport of both secreted and transmembrane proteins into the endoplasmic reticulum in yeast is promoted by the same protein translocation complex, including the Sec61, Sec62 and Sec63 proteins. All the secreted proteins possess a signal sequence at their N-termini which is recognized by the translocation complex and serves as an ER targeting signal. A typical signal sequence is comprised of several positively charged residues at the N-terminus followed by a hydrophobic core and a C-terminal site of processing by signal peptidase. Some transmembrane proteins, for example, metabotropic glutamate receptors and vasoactive intestinal polypeptide receptors, also possess the N-terminal signal sequences, whereas some do not. In the latter case, a first transmembrane domain is believed to interact with the ER translocation machinery. The use of the α -factor leader sequence may, therefore, be particularly desirable for functional expression of certain receptors.

In certain embodiments, it will be desirable to further modify the yeast cells of the present invention. For example, in one embodiment it will be desirable to disrupt the yeast calnexin-like gene, *CNE1*, to improve receptor transport from the endoplasmic reticulum to the Golgi. In yet other embodiments, it will be desirable to overexpress the gene encoding Ast1, to increase transport from the Golgi to the plasma membrane. In yet other embodiments, it will be desirable to disrupt END3 and/or END4, to prevent receptor internalization.

In preferred embodiments, certain heterologous receptors are expressed using a leader sequence other than an α factor leader sequence.

35 *IX. Test Compounds*

Exogenously added compounds

A recent trend in medicinal chemistry includes the production of mixtures of compounds, referred to as libraries. While the use of libraries of peptides is well established in the art, new techniques have been developed which have allowed the production of mixtures of other compounds, such as benzodiazepines (Bunin et al. 1992. *J. Am. Chem. Soc.* 114:10987; DeWitt et al. 1993. *Proc. Natl. Acad. Sci. USA* 90:6909)

5 peptoids (Zuckermann. 1994. J. Med. Chem. 37:2678) oligocarbamates (Cho et al. 1993. Science. 261:1303), and hydantoins (DeWitt et al. supra). Rebek et al. have described an approach for the synthesis of molecular libraries of small organic molecules with a diversity of 104-105 (Carell et al. 1994. Angew. Chem. Int. Ed. Engl. 33:2059; Carell et al. Angew. Chem. Int. Ed. Engl. 1994. 33:2061).

10 The compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries, synthetic library methods requiring deconvolution, the 'one-bead one-compound' library method, and synthetic library methods using affinity chromatography selection. The
15 biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. Anticancer Drug Des. 1997. 12:145).

 In one embodiment, the test compound is a peptide or peptidomimetic. In another, preferred embodiment, the compounds are small, organic non-peptidic
20 compounds.

 Other exemplary methods for the synthesis of molecular libraries can be found in the art, for example in: Erb et al. 1994. Proc. Natl. Acad. Sci. USA 91:11422; Horwell et al. 1996 Immunopharmacology 33:68; and in Gallop et al. 1994. J. Med. Chem. 37:1233. In addition, libraries such as those described in the commonly owned
25 applications U.S.S.N. 08/864,241, U.S.S.N. 08/864,240 and U.S.S.N. 08/835,623 can be used to provide compounds for testing in the present invention. The contents of each of these applications is expressly incorporated herein by this reference.

 Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP
30 '409), plasmids (Cull et al. (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla et al. (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

35 In certain embodiments, the test compounds are exogenously added to the yeast cells expressing a recombinant receptor and compounds that modulate signal transduction via the receptor are selected. In other embodiments, the yeast cells express the compounds to be tested. For example, a culture of the subject yeast cells can be further modified to collectively express a peptide library as described in more detail in
40 PCT Publication WO 94/23025 the contents of which is expressly incorporated herein by this reference.

5 Other types of peptide libraries may also be expressed, see, for example, U.S. Patents 5,270,181 and 5,292,646; and PCT publication WO94/ 02502). In still another embodiment, the combinatorial polypeptides are produced from a cDNA library.

Exemplary compounds which can be screened for activity include, but are not limited to, peptides, nucleic acids, carbohydrates, small organic molecules, and natural
10 product extract libraries. In such embodiments, both compounds which agonize or antagonize the receptor- or channel-mediated signaling function can be selected and identified.

Autocrine Yeast Cells

15 In certain embodiments, yeast cells can be engineered to produce the compounds to be tested. This assay system has the advantage of increasing the effective concentration of the compound to be tested. In one embodiment, a method such as that described in WO 94/23025 can be utilized.

Other methods can also be used. For example, peptide libraries are systems
20 which simultaneously display, in a form which permits interaction with a target, a highly diverse and numerous collection of peptides. These peptides may be presented in solution (Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull et al. (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin
25 (1990) *Science* 249:404-406); (Cwirla et al. (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*). Many of these systems are limited in terms of the maximum length of the peptide or the composition of the peptide (e.g., Cys excluded). Steric factors, such as the proximity of a support, may
30 interfere with binding. Usually, the screening is for binding in vitro to an artificially presented target, not for activation or inhibition of a cellular signal transduction pathway in a living cell. While a cell surface receptor may be used as a target, the screening will not reveal whether the binding of the peptide caused an allosteric change in the conformation of the receptor.

35 The Ladner et al. patent, USSN 5,096,815, describes a method of identifying novel proteins or polypeptides with a desired DNA binding activity. Semi-random ("variegated") DNA encoding a large number of different potential binding proteins is introduced, in expressible form, into suitable yeast cells. The target DNA sequence is incorporated into a genetically engineered operon such that the binding of the protein or
40 polypeptide will prevent expression of a gene product that is deleterious to the gene

5 under selective conditions. Cells which survive the selective conditions are thus cells which express a protein which binds the target DNA. While it is taught that yeast cells may be used for testing, bacterial cells are preferred. The interactions between the protein and the target DNA occur only in the cell (and then only in the nucleus), not in the periplasm or cytoplasm, and the target is a nucleic acid, and not a receptor protein.

10 Substitution of random peptide sequences for functional domains in cellular proteins permits some determination of the specific sequence requirements for the accomplishment of function. Though the details of the recognition phenomena which operate in the localization of proteins within cells remain largely unknown, the constraints on sequence variation of mitochondrial targeting sequences and protein

15 secretion signal sequences have been elucidated using random peptides (Lemire et al., *J. Biol. Chem.*(1989) 264, 20206 and Kaiser et al. (1987) *Science* 235:312, respectively).

In certain embodiments of the instant invention, the compounds tested are in the form of peptides from a peptide library. The peptide library of the present invention takes the form of a cell culture, in which essentially each cell expresses one, and usually

20 only one, peptide of the library. While the diversity of the library is maximized if each cell produces a peptide of a different sequence, it is usually prudent to construct the library so there is some redundancy. Depending on size, the combinatorial peptides of the library can be expressed as is, or can be incorporated into larger fusion proteins. The fusion protein can provide, for example, stability against degradation or denaturation, as

25 well as a secretion signal if secreted. In an exemplary embodiment of a library for intracellular expression, e.g., for use in conjunction with intracellular target receptors, the polypeptide library is expressed as thioredoxin fusion proteins (see, for example, U.S. Patents 5,270,181 and 5,292,646; and PCT publication WO94/ 02502). The combinatorial peptide can be attached one the terminus of the thioredoxin protein, or, for

30 short peptide libraries, inserted into the so-called active loop.

In one embodiment, the peptide library is derived to express a combinatorial library of polypeptides which are not based on any known sequence, nor derived from cDNA. That is, the sequences of the library are largely random. In preferred

35 embodiments, the combinatorial polypeptides are in the range of 3-100 amino acids in length, more preferably at least 5-50, and even more preferably at least 10, 13, 15, 20 or 25 amino acid residues in length. Preferably, the polypeptides of the library are of uniform length. It will be understood that the length of the combinatorial peptide does not reflect any extraneous sequences which may be present in order to facilitate expression, e.g., such as signal sequences or invariant portions of a fusion protein.

40 In another embodiment, the peptide library is a combinatorial library of polypeptides which are based at least in part on a known polypeptide sequence or a

5 portion thereof (not a cDNA library). That is, the sequences of the library is semi-random, being derived by combinatorial mutagenesis of a known sequence. See, for example, Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) *J. Biol. Chem.* 267:16007-16010; Griffiths et al. (1993) *EMBO J* 12:725-734; Clackson et al. (1991) *Nature* 352:624-628; and Barbas et al. (1992) *PNAS* 89:4457-4461. Accordingly, polypeptide(s) which are known ligands
10 for a target receptor can be mutagenized by standard techniques to derive a variegated library of polypeptide sequences which can further be screened for agonists and/or antagonists. For example, the surrogate ligand can be mutagenized to generate a library of peptides with some relationship to the original peptide. This library can be expressed
15 in a reagent cell of the present invention, and other receptor activators can be isolated from the library. This may permit the identification of even more potent surrogate ligands.

In still another embodiment, the combinatorial polypeptides are produced from a cDNA library.

20 In a preferred embodiment of the present invention, the yeast cells collectively produce a "peptide library", preferably including at least 10^3 to 10^7 different peptides, so that diverse peptides may be simultaneously assayed for the ability to interact with the exogenous receptor. In an especially preferred embodiment, at least some peptides of the peptide library are secreted into the periplasm, where they may interact with the
25 "extracellular" binding site(s) of an exogenous receptor. They thus mimic more closely the clinical interaction of drugs with cellular receptors. This embodiment optionally may be further improved (in assays not requiring pheromone secretion) by preventing pheromone secretion, and thereby avoiding competition between the peptide and the pheromone for signal peptidase and other components of the secretion system.

30 In certain embodiments of the present invention, the peptides of the library are encoded by a mixture of DNA molecules of different sequence. Each peptide-encoding DNA molecule is ligated with a vector DNA molecule and the resulting recombinant DNA molecule is introduced into a yeast cell. Since it is a matter of chance which peptide -encoding DNA molecule is introduced into a particular cell, it is not predictable
35 which peptide that cell will produce. However, based on a knowledge of the manner in which the mixture was prepared, one may make certain statistical predictions about the mixture of peptides in the peptide library.

The peptides of the library can be composed of constant and variable residues. If the *n*th residue is the same for all peptides of the library, it is said to be constant. If the
40 *n*th residue varies, depending on the peptide in question, the residue is a variable one. The peptides of the library will have at least one, and usually more than one, variable

5 residue. A variable residue may vary among any of two to all twenty of the genetically encoded amino acids; the variable residues of the peptide may vary in the same or different manner. Moreover, the frequency of occurrence of the allowed amino acids at a particular residue position may be the same or different. The peptide may also have one or more constant residues.

10 There are two principal ways in which to prepare the required DNA mixture. In one method, the DNAs are synthesized a base at a time. When variation is desired, at a base position dictated by the Genetic Code, a suitable mixture of nucleotides is reacted with the nascent DNA, rather than the pure nucleotide reagent of conventional polynucleotide synthesis.

15 The second method provides more exact control over the amino acid variation. First, trinucleotide reagents are prepared, each trinucleotide being a codon of one (and only one) of the amino acids to be featured in the peptide library. When a particular variable residue is to be synthesized, a mixture is made of the appropriate trinucleotides and reacted with the nascent DNA. Once the necessary "degenerate" DNA is complete,
20 it must be joined with the DNA sequences necessary to assure the expression of the peptide, as discussed in more detail below, and the complete DNA construct must be introduced into the yeast cell.

In embodiments in which the test compounds it may be desirable to express such peptides in the context of a leader sequence. Yeast cells are bounded by a lipid bilayer
25 called the plasma membrane. Between this plasma membrane and the cell wall is the periplasmic space. Peptides secreted by yeast cells cross the plasma membrane through a variety of mechanisms and thereby enter the periplasmic space. The secreted peptides are then free to interact with other molecules that are present in the periplasm or displayed on the outer surface of the plasma membrane. The peptides then either
30 undergo re-uptake into the cell, diffuse through the cell wall into the medium, or become degraded within the periplasmic space.

The test polypeptide library may be secreted into the periplasm by any of a number of exemplary mechanisms, depending on the nature of the expression system to which they are linked. In one embodiment, the peptide may be structurally linked to a
35 yeast signal sequence, such as that present in the α -factor precursor, which directs secretion through the endoplasmic reticulum and Golgi apparatus. Since this is the same route that the receptor protein follows in its journey to the plasma membrane, opportunity exists in cells expressing both the receptor and the peptide library for a specific peptide to interact with the receptor during transit through the secretory
40 pathway. This has been postulated to occur in mammalian cells exhibiting autocrine activation. Such interaction could yield activation of the response pathway during

5 transit, which would still allow identification of those cells expressing a peptide agonist. For situations in which peptide antagonists to externally applied receptor agonist are sought, this system would still be effective, since both the peptide antagonist and receptor would be delivered to the outside of the cell in concert. Thus, those cells producing an antagonist would be selectable, since the peptide antagonist would be
 10 properly and timely situated to prevent the receptor from being stimulated by the externally applied agonist.

An alternative mechanism for delivering peptides to the periplasmic space is to use the ATP-dependent transporters of the STE6/MDR1 class. This transport pathway and the signals that direct a protein or peptide to this pathway are not as well
 15 characterized as is the endoplasmic reticulum-based secretory pathway. Nonetheless, these transporters apparently can efficiently export certain peptides directly across the plasma membrane, without the peptides having to transit the ER/Golgi pathway. It is anticipated that at least a subset of peptides can be secreted through this pathway by expressing the library in context of the a-factor prosequence and terminal tetrapeptide.
 20 The possible advantage of this system is that the receptor and peptide do not come into contact until both are delivered to the external surface of the cell. Thus, this system strictly mimics the situation of an agonist or antagonist that is normally delivered from outside the cell. Use of either of the described pathways is within the scope of the invention. The present invention does not require periplasmic secretion, or, if such
 25 secretion is provided, any particular secretion signal or transport pathway.

X. Screening and Selection: Assays of Second Messenger Generation

When screening for bioactivity of compounds, intracellular second messenger generation can be measured directly. A variety of intracellular effectors have been
 30 identified as being G-protein-regulated, including adenylyl cyclase, cyclic GMP, phosphodiesterases, phosphoinositidase C, and phospholipase A₂. In addition, G proteins interact with a range of ion channels and are able to inhibit certain voltage-sensitive Ca⁺⁺ transients, as well as stimulating cardiac K⁺ channels.

In one embodiment, the GTPase enzymatic activity by G proteins can be
 35 measured in plasma membrane preparations by determining the breakdown of $\gamma^{32}\text{P}$ GTP using techniques that are known in the art (For example, see *Signal Transduction: A Practical Approach*. G. Milligan, Ed. Oxford University Press, Oxford England). When receptors that modulate cAMP are tested, it will be possible to use standard techniques for cAMP detection, such as competitive assays which quantitate [³H]cAMP
 40 in the presence of unlabelled cAMP.

5 Certain receptors stimulate the activity of phospholipase C which stimulates the breakdown of phosphatidylinositol 4,5, bisphosphate to 1,4,5-IP₃ (which mobilizes intracellular Ca⁺⁺) and diacylglycerol (DAG) (which activates protein kinase C). Inositol lipids can be extracted and analyzed using standard lipid extraction techniques. DAG can also be measured using thin-layer chromatography. Water soluble derivatives
10 of all three inositol lipids (IP₁, IP₂, IP₃) can also be quantitated using radiolabelling techniques or HPLC.

The mobilization of intracellular calcium or the influx of calcium from outside the cell can be measured using standard techniques. The choice of the appropriate calcium indicator, fluorescent, bioluminescent, metallochromic, or Ca⁺⁺-sensitive
15 microelectrodes depends on the cell type and the magnitude and time constant of the event under study (Borle (1990) *Environ Health Perspect* 84:45-56). As an exemplary method of Ca⁺⁺ detection, cells could be loaded with the Ca⁺⁺-sensitive fluorescent dye fura-2 or indo-1, using standard methods, and any change in Ca⁺⁺ measured using a fluorometer.

20 The other product of PIP₂ breakdown, DAG can also be produced from phosphatidyl choline. The breakdown of this phospholipid in response to receptor-mediated signaling can also be measured using a variety of radiolabelling techniques.

The activation of phospholipase A₂ can easily be quantitated using known techniques, including, for example, the generation of arachadonate in the cell.

25 In the case of certain receptors, it may be desirable to screen for changes in cellular phosphorylation. Such assay formats may be useful when the receptor of interest is a receptor tyrosine kinase. For example, yeast transformed with the FGF receptor and a ligand which binds the FGF receptor could be screened using colony immunoblotting (Lyons and Nelson (1984) *Proc. Natl. Acad. Sci. USA* 81:7426-7430)
30 using anti-phosphotyrosine. In addition, tests for phosphorylation could be useful when a receptor which may not itself be a tyrosine kinase, activates protein kinases that function downstream in the signal transduction pathway. Likewise, it is noted that protein phosphorylation also plays a critical role in cascades that serve to amplify signals generated at the receptor. Multi-kinase cascades allow not only signal
35 amplification but also signal divergence to multiple effectors that are often cell-type specific, allowing a growth factor to stimulate mitosis of one cell and differentiation of another.

One such cascade is the MAP kinase pathway that appears to mediate both mitogenic, differentiation and stress responses in different cell types. Stimulation of
40 growth factor receptors results in Ras activation followed by the sequential activation of c-Raf, MEK, and p44 and p42 MAP kinases (ERK1 and ERK2). Activated MAP kinase

5 then phosphorylates many key regulatory proteins, including p90RSK and Elk-1 that are phosphorylated when MAP kinase translocates to the nucleus. Homologous pathways exist in mammalian and yeast cells. For instance, an essential part of the *S. cerevisiae* pheromone signaling pathway is comprised of a protein kinase cascade composed of the products of the STE11, STE7, and FUS3/KSS1 genes (the latter pair are distinct and
10 functionally redundant). Accordingly, phosphorylation and/or activation of members of this kinase cascade can be detected and used to quantitate receptor engagement. Phosphotyrosine specific antibodies are available to measure increases in tyrosine phosphorylation and phospho-specific antibodies are commercially available (New England Biolabs, Beverly, MA).

15 Modified methods for detecting receptor-mediated signal transduction exist and one of skill in the art will recognize suitable methods that may be used to substitute for the example methods listed.

In one embodiment, the indicator gene can be used for detection. In one embodiment an indicator gene is an unmodified endogenous gene. For example, the
20 instant method can rely on detecting the transcriptional level of such pheromone system pathway responsive endogenous genes as the *Bar1* or *Fus1*, *Fus 2*, mating factor, *Ste3*, *Ste13*, *Kex1*, *Ste2*, *Ste6*, *Ste7*, *sSst2*, or *Chs1*. (Appeltauer and Zchstetter. 1989. Eur. J. Biochem. 181:243)

In other embodiments, the sensitivity of an endogenous indicator gene can be
25 enhanced by manipulating the promoter sequence at the natural locus for the indicator gene. Such manipulation may range from point mutations to the endogenous regulatory elements to gross replacement of all or substantial portions of the regulatory elements. The previous discussion of mutations with regard to G proteins and G protein coupled receptors is reiterated here.

30 For example, in the case of the *Bar1* gene, the promoter of the gene can be modified to enhance the transcription of *Bar1* upon activation of the yeast pheromone system pathway. *Bar1* gene transcription is iactivated upon exposure of yeast cells to mating factor. The sequence of the *Bar1* gene is known in the art (see e.g., U.S. patent 4,613,572). Moreover, the sequences required for α -factor-enhanced expression of the
35 *Bar1*, and other pheromone responsive genes have been identified. (Appeltauer and Achstetter 1989. Eur. J. Biochem. 181:243; Hagen et al. 1991. Mol. Cell. Biol. 11:2952). In an exemplary embodiment, the yeast *Bar1* promoter can be engineered by mutagenesis to be more responsive, e.g., to more strongly promoter gene transcription, upon stimulation of the yeast pheromone pathway. Standard techniques for
40 mutagenizing the promoter can be used. In such embodiments, it is desirable that the conserved oligonucleotide motif described by Appeltaure et al. be conserved.

In other embodiments, the modulation of a receptor by a test compound can result in a change in the transcription of a gene, which is not normally pheromone responsive. In preferred embodiments, the gene is easily detectable. For example, in a preferred embodiment, the subject assay can be used to measure Pho5, a secreted acid phosphatase. Acid phosphatase activity can be measured using standard techniques.

Many reporter genes and transcriptional regulatory elements are known to those of skill in the art and others may be identified or synthesized by methods known to those of skill in the art. Reporter genes include any gene that expresses a detectable gene product, which may be RNA or protein. Preferred reporter genes are those that are readily detectable. The reporter gene may also be included in the construct in the form of a fusion gene with a gene that includes desired transcriptional regulatory sequences or exhibits other desirable properties.

Examples of reporter genes include, but are not limited to CAT (chloramphenicol acetyl transferase) (Alton and Vapnek (1979), Nature 282: 864-869) luciferase, and other enzyme detection systems, such as beta-galactosidase; firefly luciferase (deWet et al. (1987), Mol. Cell. Biol. 7:725-737); bacterial luciferase (Engebrecht and Silverman (1984), PNAS 1: 4154-4158; Baldwin et al. (1984), Biochemistry 23: 3663-3667); alkaline phosphatase (Toh et al. (1989) Eur. J. Biochem. 182: 231-238, Hall et al. (1983) J. Mol. Appl. Gen. 2: 101), human placental secreted alkaline phosphatase (Cullen and Malim (1992) Methods in Enzymol. 216:362-368) and green fluorescent protein (U.S. patent 5,491,084; WO96/23898).

5 Transcriptional control elements include, but are not limited to, promoters, enhancers, and repressor and activator binding sites. Suitable transcriptional regulatory elements may be derived from the transcriptional regulatory regions of genes whose expression is rapidly induced, generally within minutes, of contact between the cell surface protein and the effector protein that modulates the activity of the cell surface

10 protein. Examples of such genes include, but are not limited to, the immediate early genes (see, Sheng et al. (1990) *Neuron* 4: 477-485), such as c-fos. Immediate early genes are genes that are rapidly induced upon binding of a ligand to a cell surface protein. The transcriptional control elements that are preferred for use in the gene constructs include transcriptional control elements from immediate early genes, elements

15 derived from other genes that exhibit some or all of the characteristics of the immediate early genes, or synthetic elements that are constructed such that genes in operative linkage therewith exhibit such characteristics. The characteristics of preferred genes from which the transcriptional control elements are derived include, but are not limited to, low or undetectable expression in quiescent cells, rapid induction at the

20 transcriptional level within minutes of extracellular stimulation, induction that is transient and independent of new protein synthesis, subsequent shut-off of transcription requires new protein synthesis, and mRNAs transcribed from these genes have a short half-life. It is not necessary for all of these properties to be present.

Other promoters and transcriptional control elements, in addition to those

25 described above, include the vasoactive intestinal peptide (VIP) gene promoter (cAMP responsive; Fink et al. (1988), *Proc. Natl. Acad. Sci.* 85:6662-6666); the somatostatin gene promoter (cAMP responsive; Montminy et al. (1986), *Proc. Natl. Acad. Sci.* 83:6682-6686); the proenkephalin promoter (responsive to cAMP, nicotinic agonists, and phorbol esters; Comb et al. (1986), *Nature* 323:353-356); the phosphoenolpyruvate

30 carboxy-kinase gene promoter (cAMP responsive; Short et al. (1986), *J. Biol. Chem.* 261:9721-9726); the NGFI-A gene promoter (responsive to NGF, cAMP, and serum; Changelian et al. (1989), *Proc. Natl. Acad. Sci.* 86:377-381); and others that may be known to or prepared by those of skill in the art.

In certain assays it may be desirable to use changes in growth in the screening

35 procedure. For example, one of the consequences of activation of the pheromone signal pathway in wild-type yeast is growth arrest. If one is testing for an antagonist of a G protein-coupled receptor, this normal response of growth arrest can be used to select cells in which the pheromone response pathway is inhibited. That is, cells exposed to both a known agonist and a peptide of unknown activity will be growth arrested if the

40 peptide is neutral or an agonist, but will grow normally if the peptide is an antagonist. Thus, the growth arrest response can be used to advantage to discover peptides that function as antagonists.

5 In certain embodiments, when searching for compounds which can function as agonists of G protein-coupled receptors, or other pheromone system proteins, the growth arrest consequent to activation of the pheromone response pathway is an undesirable effect since cells that bind agonists stop growing while surrounding cells that fail to bind agonists will continue to grow. The cells of interest, then, will be overgrown or their
 10 detection obscured by the background cells, confounding identification of the cells of interest. To overcome this problem the present invention teaches engineering the cell such that: 1) growth arrest does not occur as a result of exogenous signal pathway activation (e.g., by inactivating the FAR1 gene); and/or 2) a selective growth advantage is conferred by activating the pathway (e.g., by transforming an auxotrophic mutant with
 15 a HIS3 gene under the control of a pheromone-responsive promoter, and applying selective conditions).

Alternatively, the promoter may be one which is repressed by the receptor pathway, thereby preventing expression of a product which is deleterious to the cell. With a receptor repressed promoter, one screens for agonists by linking the promoter to
 20 a deleterious gene, and for antagonists, by linking it to a beneficial gene. Repression may be achieved by operably linking a receptor-induced promoter to a gene encoding mRNA which is antisense to at least a portion of the mRNA encoded by the marker gene (whether in the coding or flanking regions), so as to inhibit translation of that mRNA. Repression may also be obtained by linking a receptor-induced promoter to a gene
 25 encoding a DNA-binding repressor protein, and incorporating a suitable operator site into the promoter or other suitable region of the marker gene.

In the case of yeast, exemplary positively selectable (beneficial) genes include the following: *URA3*, *LYS2*, *HIS3*, *LEU2*, *TRP1*; *ADE1*, 2, 3, 4, 5, 7, 8; *ARG1*, 3, 4, 5, 6, 8; *HIS1*, 4, 5; *ILV1*, 2, 5; *THR1*, 4; *TRP2*, 3, 4, 5; *LEU1*, 4; *MET2*, 3, 4, 8, 9, 14, 16, 19;
 30 *URA1*, 2, 4, 5, 10; *HOM3*, 6; *ASP3*; *CHO1*; *ARO* 2, 7; *CYS3*; *OLE1*; *INO1*, 2, 4; *PRO1*, 3
 Countless other genes are potential selective markers. The above are involved in well-characterized biosynthetic pathways. The imidazoleglycerol phosphate dehydratase (IGP dehydratase) gene (*HIS3*) is preferred because it is both quite sensitive and can be selected over a broad range of expression levels. In the simplest case, the cell is
 35 auxotrophic for histidine (requires histidine for growth) in the absence of activation. Activation leads to synthesis of the enzyme and the cell becomes prototrophic for histidine (does not require histidine). Thus the selection is for growth in the absence of histidine. Since only a few molecules per cell of IGP dehydratase are required for histidine prototrophy, the assay is very sensitive.

40 In another version of the assay, cells can be selected for resistance to aminotriazole (AT), a drug that inhibits the activity of IGP dehydratase. Cells with low,

5 fixed level of expression of HIS3 are sensitive to the drug, while cells with higher levels are resistant. The amount of AT can be selected to inhibit cells with a basal level of HIS3 expression (whatever that level is) but allow growth of cells with an induced level of expression. In this case selection is for growth in the absence of histidine and in the presence of a suitable level of AT.

10 In appropriate assays, so-called counterselectable or negatively selectable genes may be used. Suitable genes include: URA3 (orotidine-5'-phosphate decarboxylase; inhibits growth on 5-fluoroorotic acid), LYS2 (2-amino adipate reductase; inhibits growth on α -amino adipate as sole nitrogen source), CYH2 (encodes ribosomal protein L29; cycloheximide-sensitive allele is dominant to resistant allele), CAN1 (encodes
15 arginine permease; null allele confers resistance to the arginine analog canavanin), and other recessive drug-resistant markers.

In one example, the reporter gene effects yeast cell growth. The natural response to signal transduction via the yeast pheromone system response pathway is for cells to undergo growth arrest. This is a preferred way to select for antagonists of a
20 ligand/receptor pair that stimulates a the pathway. An antagonist would inhibit the activation of the pathway; hence, the cell would be able to grow. Thus, the FAR1 gene may be considered an endogenous counterselectable marker. The FAR1 gene is preferably inactivated when screening for agonist activity.

The reporter gene may also be a screenable gene. The screened characteristic
25 may be a change in cell morphology, metabolism or other screenable features. Suitable markers include beta-galactosidase (Xgal, C₁₂FDG, Salmon-gal, Magenta-Gal (latter two from Biosynth Ag)), alkaline phosphatase, horseradish peroxidase, exo-glucanase (product of yeast exb1 gene; nonessential, secreted); luciferase; bacterial green fluorescent protein; (human placental) secreted alkaline phosphatase (SEAP); and
30 chloramphenicol transferase (CAT). Some of the above can be engineered so that they are secreted (although not β -galactosidase). A preferred screenable marker gene is beta-galactosidase; yeast cells expressing the enzyme convert the colorless substrate Xgal into a blue pigment. Again, the promoter may be receptor-induced or receptor-inhibited.

35 XI. Other Optional Alterations to Yeast Cells

The choice of appropriate host cell will also be influenced by the choice of detection signal. For instance, reporter constructs can provide a selectable or screenable trait upon transcriptional activation (or inactivation) in response to a signal transduction pathway coupled to the target receptor. The indicator gene may be an unmodified gene
40 already in the host cell pathway, such as the genes responsible for growth arrest in yeast.

5 In certain embodiments a host cell gene may be operably linked to a "receptor-responsive" promoter. Alternatively, it may be a heterologous gene that has been so linked. Suitable genes and promoters are discussed below.

To achieve optimal selection or screening, the host cell phenotype will be considered. For example, introducing a pheromone-responsive chimeric HIS3 gene into
 10 a yeast that has a wild-type HIS3 gene would frustrate genetic selection. Thus, to achieve nutritional selection, an auxotrophic strain is preferred. Yeast strains that are auxotrophic for histidine (HIS3) are known, see Struhl and Hill, (1987) *Mol. Cell. Biol.*, 7:104; Fasullo and Davis, *Mol. Cell. Biol.*, (1988) 8:4370. The HIS3 (imidazoleglycerol phosphate dehydratase) gene has been used as a selective marker in yeast. See Sikorski
 15 and Heiter, (1989) *Genetics*, 122:19; Struhl, et al., P.N.A.S. (1979) 76:1035; and, for FUS1-HIS3 fusions, see Stevenson, et al., (1992) *Genes Dev.*, 6:1293.

In certain embodiments, the host yeast cell can be modified in other ways. For example, it may be desirable to inactivate, such as by mutation or deletion, a homologous receptor, e.g., a pheromone receptor, present in the cell in order to
 20 minimize interference with signaling via the heterologous receptor. "Inactivation", with respect to genes of the host cell, means that production of a functional gene product is prevented or inhibited. Inactivation may be achieved by deletion of the gene, mutation of the promoter so that expression does not occur, or mutation of the coding sequence so that the gene product is inactive. Inactivation may be partial or total.

25 In a preferred embodiment of the subject assay, the yeast cells possess one or more of the following characteristics: (a) the endogenous FUS1 gene has been inactivated; (b) the endogenous SST2 gene, and/or other genes involved in desensitization, have been inactivated; (c) if there is a homologous, endogenous receptor gene it has been inactivated; and (d) if the yeast produces an endogenous ligand
 30 to the exogenous receptor, the genes encoding for the ligand been inactivated.

It is desirable that the exogenous receptor be exposed on a continuing basis to the peptides. In some instances, this may result in desensitization of the pheromone pathway to the stimulus. For example, the mating signal transduction pathway is known to become desensitized by several mechanisms including pheromone degradation and
 35 modification of the function of the receptor, G proteins and/or downstream elements of the pheromone signal transduction by the products of the SST2, STE50, AFR1 (Konopka, J.B. (1993) *Mol. Cell. Biol.* 13:6876-6888) and SGV1, MSG5, and SIG1 genes. Selected mutations in these genes can lead to hypersensitivity to pheromone and an inability to adapt to the presence of pheromone. For example, introduction of
 40 mutations that interfere with function into strains expressing heterologous G protein-coupled receptors constitutes a significant improvement on wild type strains and enables

5 the development of extremely sensitive bioassays for compounds that interact with the receptors. Other mutations e.g. STE50, *sgv1*, *bar1*, *ste2*, *ste3*, *pik1*, *msg5*, *sig1*, and *aft1*, have the similar effect of increasing the sensitivity of the bioassay. Thus desensitization may be avoided by mutating (which may include deleting) the SST2 gene so that it no longer produces a functional protein, or by mutating one of the other genes listed above.

10 In certain embodiments, it will be desirable to complement the host yeast cells, e.g., least partial function of an inactivated gene of the host cell can be supplied by an exogenous nucleic acid. For instance, yeast cells can be "mammalianized", and even "humanized", by complementation of receptor and signal transduction proteins with mammalian homologues. To illustrate, inactivation of a yeast Byr2/Ste11 gene can be
15 complemented by expression of a human MEKK gene.

Complementations for use in the subject assay can be constructed without any undue experimentation. Indeed, many yeast genetic complementations with mammalian signal transduction proteins have been described in the art. For example, Mosteller et al. (1994) Mol Cell Biol 14:1104-12 demonstrates that human Ras proteins can
20 complement loss of ras mutations in *S. cerevisiae*. Moreover, Toda et al. (1986) Princess Takamatsu Symp 17: 253-60 have shown that human ras proteins can complement the loss of RAS1 and RAS2 proteins in yeast, and hence are functionally homologous. Both human and yeast RAS proteins can stimulate the magnesium and guanine nucleotide-dependent adenylate cyclase activity present in yeast membranes.
25 Ballester et al. (1989) Cell 59: 681-6 describe a vector to express the mammalian GAP protein in the yeast *S. cerevisiae*. When expressed in yeast, GAP inhibits the function of the human ras protein, and complements the loss of IRA1. IRA1 is a yeast gene that encodes a protein with homology to GAP and acts upstream of RAS. Mammalian GAP can therefore function in yeast and interact with yeast RAS. Wei et al. (1994) Gene 151:
30 279-84 describes that a human Ras-specific guanine nucleotide-exchange factor, Cdc25GEF, can complement the loss of CDC25 function in *S. cerevisiae*. Martegani et al. (1992) EMBO J 11: 2151-7 describe the cloning by functional complementation of a mouse cDNA encoding a homolog of CDC25, a *Saccharomyces cerevisiae* RAS activator. Vojtek et al. (1993) J Cell Sci 105: 777-85 and Matviw et al. (1992) Mol Cell
35 Biol 12: 5033-40 describe how a mouse CAP protein, e.g., an adenylyl cyclase associated protein associated with ras-mediated signal transduction, can complements defects in *S. cerevisiae*. Papasavvas et al. (1992) Biochem Biophys Res Commun 184:1378-85 also suggest that inactivated yeast adenylyl cyclase can be complemented by a mammalian adenylyl cyclase gene. Hughes et al. (1993) Nature 364: 349-52 describe
40 the complementation of *byr1* in fission yeast by mammalian MAP kinase kinase (MEK). Parissenti et al. (1993) Mol Cell Endocrinol 98: 9-16 describes the reconstitution of bovine protein kinase C (PKC) in yeast. The Ca(2+) and phospholipid-dependent

5 Ser/Thr kinase PKC plays important roles in the transduction of cellular signals in mammalian cells. Marcus et al. (1995) PNAS 92: 6180-4 suggests the complementation of *shk1* null mutations in *S. pombe* by the either the structurally related *S. cerevisiae* Ste20 or mammalian p65PAK protein kinases.

10 XII. Pharmaceutical Preparations of Identified Compounds

After identifying certain test compounds as potential surrogate ligands, or receptor antagonists, the practitioner of the subject assay will continue to test the efficacy and specificity of the selected compounds both *in vitro* and *in vivo*. Whether for subsequent *in vivo* testing, or for administration to an animal as an approved drug, agents
15 identified in the subject assay can be formulated in pharmaceutical preparations for *in vivo* administration to an animal, preferably a human.

The compounds selected in the subject assay, or a pharmaceutically acceptable salt thereof, may accordingly be formulated for administration with a biologically acceptable medium, such as water, buffered saline, polyol (for example, glycerol,
20 propylene glycol, liquid polyethylene glycol and the like) or suitable mixtures thereof. The optimum concentration of the active ingredient(s) in the chosen medium can be determined empirically, according to procedures well known to medicinal chemists. As used herein, "biologically acceptable medium" includes any and all solvents, dispersion media, and the like which may be appropriate for the desired route of administration of the pharmaceutical preparation. The use of such media for pharmaceutically active
25 substances is known in the art. Except insofar as any conventional media or agent is incompatible with the activity of the compound, its use in the pharmaceutical preparation of the invention is contemplated. Suitable vehicles and their formulation inclusive of other proteins are described, for example, in the book *Remington's Pharmaceutical Sciences* (Remington's
30 Pharmaceutical Sciences. Mack Publishing Company, Easton, Pa., USA 1985). These vehicles include injectable "deposit formulations". Based on the above, such pharmaceutical formulations include, although not exclusively, solutions or freeze-dried powders of the compound in association with one or more pharmaceutically acceptable
35 vehicles or diluents, and contained in buffered media at a suitable pH and isosmotic with physiological fluids. In preferred embodiment, the compound can be disposed in a sterile preparation for topical and/or systemic administration. In the case of freeze-dried preparations, supporting excipients such as, but not exclusively, mannitol or glycine may be used and appropriate buffered solutions of the desired volume will be provided so as
40 to obtain adequate isotonic buffered solutions of the desired pH. Similar solutions may also be used for the pharmaceutical compositions of compounds in isotonic solutions of

5 the desired volume and include, but not exclusively, the use of buffered saline solutions with phosphate or citrate at suitable concentrations so as to obtain at all times isotonic pharmaceutical preparations of the desired pH, (for example, neutral pH).

10 All patents, published patent applications and other references disclosed herein are hereby expressly incorporated by reference. In addition the contents of USSNs: 08/322,137; 08/463, 181 and 08/946,298 are hereby incorporated by this reference.

EXEMPLIFICATION

15 The invention now being generally described will be more readily understood by reference to the following examples, which are included merely for purposes of illustration of certain aspects and embodiments of the present invention and are not intended to limit the invention.

20 Example 1: Amplification of the yeast pheromone response pathway using Ste5

Yeast strains have been constructed that respond to stimulation of heterologously expressed GPCRs by overexpressing Ste5p (either the wildtype form or hypersensitive forms), using constructs that drive expression of Ste5p from a pheromone responsive promoter (*e.g.*, FUS1), thereby amplifying the activity of the pheromone response pathway (PRP) upon ligand stimulation of the GPCR.

To express STE5p from the FUS1 promoter, the STE5 open reading frame was PCR-amplified from genomic yeast DNA using oligonucleotides o440 (5'-
30 GCTAAGATAGAGTTTGCTCCGGACGG-3') (SEQ ID NO: 1) and o443 (5'-CCGGATCCATGATGGAACTCCTACAGAC-3') (SEQ ID NO: 2) to PCR amplify the 5' half of the open reading frame and o441 (5'-ATACATTTTACATTACCCCCTGG-3') (SEQ ID NO: 3) and o444 (5'-GGGCGGCCGCTATATATAATCCATATGGAGG-3') (SEQ ID NO: 4) to amplify the 3' half of the open reading frame.

35 The 5' product was digested with BamHI and EcoRI, the 3' product was digested with EcoRI and EagI, and the PCR products were used to replace the *lacZ* open reading frame in plasmids Cp1944 (genotype *URA3 CEN6 ARSH4 bla FUS1p-lacZ*) and Cp1947 (genotype *URA3 2m -ori REP3 AmpR FUS1p-lacZ*). Specifically, Cp1944 and Cp1947 were each digested with BamH I and Eag I, the 6.0 kb (from Cp1944) and 6.8 kb (from
40 Cp1947) fragments were isolated, and these vector fragments were combined with the 5' and 3' PCR products in a three-fragment ligation.

5 Each of the resulting recombinant plasmids, Cp5373 (genotype *URA3 CEN6 ARSH4 bla FUS1p-STE5*) and Cp5348 (genotype *URA3 2m-ori REP3 bla FUS1p-STE5*), expresses a 921 residue protein consisting of the entire 917 residues of STE5p with a short, vector-derived leader sequence MetAlaGlySer (SEQ ID NO: 5). It should be noted that the STE5 open reading frame used in this experiment differs from the
 10 publicly deposited sequence in Genbank (Accession number Z47746) at residue 877 of the Genbank sequence. The Ste 5 PCR product used herein encodes a glycine at this position rather than the expected aspartate. Since this difference was observed in multiple independent PCR products, the difference likely reflects a polymorphism among yeast strains. Indeed, when this glycine-encoding triplet was intentionally
 15 mutagenized to encode aspartate, no difference was observed for any *STE5*-dependent phenotype tested.

To express the hyperactive alleles of STE5 (either STE5(T52M) or STE5(S18R,T52M)) from the FUS1 promoter, an approximately 300 bp PCR fragment
 20 carrying the mutated codons was used to replace the corresponding wild-type sequence in Cp5373. Specifically, oligonucleotides o498 (5'-GGGGATCCATGATGGAAACT CCTACAGACAATATAG-3') (SEQ ID NO: 6) and o499 (5'-CGAACGTCT GCCTTTGAAGTGGTGCC-3') (SEQ ID NO: 7) were used to PCR-amplify the fragments carrying the T52M mutation from Cp5516 (obtained from D. Jenness; see
 25 Mol. Cell. Biol. 14:1054, 1994) and the S18R,T52M mutations from Cp5515 (also obtained from D. Jenness; see Mol. Cell. Biol. 14:1054, 1994).

Cp5373 was cut with BamH I and Afl II, and the approximately 8.1 kb fragment was gel-isolated for ligation with each of the PCR products, also cut with BamH I and Afl II. No significant difference were found in the phenotypes conferred by the two
 30 resulting recombinant plasmids, Cp5643 (genotype *URA3 CEN6 ARSH4 bla FUS1p-STE5(S18R,T52M)*) and Cp5644 (genotype *URA3 CEN6 REP3 bla FUS1p-STE5(T52M)*).

Yeast strains expressing heterologous GPCRs can be constructed as previously
 35 described (see e.g. WO 94/23025). In this example, to create a yeast strain expressing a heterologous GPCR and a Ste5 expression plasmid, yeast strain CY10103 (genotype MATa *ste18g6-3841 far1D1442 cyh2 tbt1-1 FUS1p-HIS3 can1 ste14::trp1::LYS2 ste3D1156 gpa1(41)-Gai2 lys2 ura3 leu2 trp1 his3 ade2D3447 ade8D3457*) was transformed by electroporation with a C5a receptor-bearing plasmid (Cp1303; genotype
 40 *LEU2 PGKp-C5aR 2m-ori REP3 bla*), a reporter plasmid (Cp1584; genotype *bla TRP1 2m-ori FUS1p-lacZ*), and plasmids that express either wild-type STE5p (Cp5373; genotype *URA3 CEN6 ARSH4 bla FUS1p-STE5*), a hyperactive STE5p mutant (Cp5643;

5 genotype *URA3 CEN6 ARSH4 bla FUS1p-STE5(S18R,T52M)*), or no STE5p (Cp1015; genotype *URA3 CEN6 ARSH4 bla*). Triple transformants were selected on SC plates lacking uracil, tryptophan, and leucine.

After overnight growth in selective media, the strains were tested for responsiveness to the ligand ChaCha (N-methyl-Phe-Lys-Pro-dCha-Cha-dArg, where
 10 Cha is cyclohexylalanine), a potent agonist of the C5a receptor, by monitoring β -galactosidase activity after 4 hours incubation of yeast with varying concentrations of ChaCha. Yeast strains bearing the C5a receptor carrying the low-copy, FUS1p-STE5-bearing plasmid showed increases in their maximal responses to the agonist ChaCha compared to strains without the FUS1p-STE5 plasmid (Figures 1A-D). Results suggest
 15 an effect of the low-copy, FUS1p-STE5 construct in enhancing the maximal signal in response to ChaCha. In Figures 1A-D, sibling strains CY13970 and 13971 overexpressing Ste5 from the FUS 1 promoter (genotype *MATa ste18g6-3841 far 1 D1442 cyh2 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3D1156 gpa1(41)-Gai2 lys2 ura3 leu2 trp1 his3 ade2D3447 ade8D3457/ URA3 CEN6 ARSH4 AmpR FUS1p-STE5/LEU2 PGKp-C5aR 2m REP3 AmpR/AmpR TRP1 2m FUS1p-lacZ*) were assayed
 20 for β galactosidase activity in triplicate, as were siblings CY13974 and 13975 carrying the wildtype Ste5 (genotype *MATa ste18g6-3841 far 1 D1442 cyh2 tbt1-1 fus1-HIS3 can1 ste14::trp1::LYS2 ste3D1156 gpa1(41)-Gai2 lys2 ura3 leu2 trp1 his3 ade2D3447 ade8D3457/ URA3 CEN6 ARSH4 AmpR /LEU2 PGKp-C5aR 2m REP3 AmpR/AmpR TRP1 2m FUS1p-lacZ*). The values for the siblings did not differ and so were averaged
 25 for graphing. The graphs show the concentration dependence of both β -galactosidase activities and induction ratios for the strains of the two genotypes.

30 **Example 2: Amplification of the yeast pheromone response pathway using Ste4**

Plasmids were also constructed that express STE4 from the FUS1 promoter. The STE4 open reading frame was PCR-amplified from genomic DNA (strain CY1) using oligonucleotides o512 (5'-GGGCGTCTCAGATCGATGGCAGCACATCAGA
 35 TGGAC-3') (SEQ ID NO: 8) and o513 (5'-GGGCGGCCGCTATTGATAACCTGGA GACC-3') (SEQ ID NO: 9). The PCR product was digested with BsmB I and Eag I for cloning into BamH I- and Eag I-digested Cp1944 (genotype *URA3 CEN6 ARSH4 bla FUS1p-lacZ*), thereby replacing the lacZ open reading frame of Cp1944 with the STE4 open reading frame. The resulting plasmid, Cp5559 (genotype *URA3 CEN6 ARSH4 bla FUS1p-STE4*), expresses a expresses a 427 residue protein consisting of the entire 423
 40 residues of STE4p with a short, vector-derived leader sequence MetAlaGlySer (SEQ ID NO: 5).

5 A TRP1-marked version of the same construct, Cp5591, was generated by subcloning the 2.4 kb Kpn I-to-Eag I fragment (containing the FUS1p-STE4 expression unit) into Kpn I- and Eag I-digested pRS414 (Gene 110:119, 1992). Sequencing of the STE4 open reading frame indicated a single base change relative to the sequence reported as Accession Number P18852 in Genbank: base 1202 is changed from T to C ,
 10 thereby resulting in Leu to Ser substitution. This likely reflects a polymorphism among yeast strains, since the C was observed this difference in multiple independent PCR products of yeast strain CY1.

A so-called haploid-lethal allele of STE4 (Mol. Cell. Biol 14: 3223-3229, 1994)
 15 was generated by site-directed mutagenesis of Cp5559 using mutagenic primers o534 (5'-CCATTAGATTCTCAAAGAGTTTTTTCCTGCGCTATTTTCG-3') (SEQ ID NO: 10) and o535 (5'-CGAAATAGCGCAGGAAAAAACTCTTTGAGAATCTAATGG-3') (SEQ ID NO: 11). The resulting plasmid, Cp5715 (genotype *URA3 CEN6 ARSH4 AmpR FUS1p-STE4(W136R, L138F)*), encodes the W136R,L138F double mutant of
 20 STE4. The same mutagenic primers were used to generate Cp5716 (genotype *TRP1 CEN6 ARSH4 AmpR FUS1p-STE4(W136R, L138F)*) from Cp5591. The effects of these plasmids on signalling through the pheromone response pathway can be tested as for the STE5-expressing constructs as described in Example 1 above.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant:

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.